ANNUAL
REPORT 2017

The cover page shows the MMD PLATFORM: a reproducible research platform tested at Idiap to detect plant diseases, by learning from user demonstration how to manipulate and inspect leaves.

CITY OF MARTIGNY

CANTON OF VALAIS

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

www.loro.ch

www.swisscom.ch

www.epfl.ch

www.theark.ch

www.ideark.ch

www.snf.ch

www.kti.admin.ch

www.haslerstiftung.ch

CONTENTS

Messages

Organization

Finances

Research

Network

Faces

Scientific inserts

→ 25

→ 37

→ 23

→ 31

→ 36

→ 26

Artificial intelligence— Idiap at the forefront of research

T he financial year 2017 is marked by further growth in the Institute's turnover largely due to the diversity of revenue sources, including North American projects and the rising reputation of the Biometrics Center. Our dynamic commitment and the excellence of our management and employees should also be underlined.

For a long time now, we have been able to count on public sector funding and we thank the political authorities and the members of the various administrations who thus permit us to pursue the fundamental research activities that enable us to secure mandates, particularly in the industrial world.

The Institute is developing—proof that the areas of activity in which it is engaged are expanding. No one can dispute the opportunities that artificial intelligence (AI) now offers, particularly if one factors in machines' ability to develop a form of thinking that enables them to achieve real autonomy. The latest industrial robots and applications in medicine are the forerunners of an era in which the entire tertiary sector is already engaged.

It is not without pride that Idiap participates in these developments. Certain conditions are required, however, if we are to ride the wave of our success. In particular, there is a compelling need for computer development engineers who are able to "set to music" the results of the research undertaken at the Institute. To sustain our activities and render them viable in the longer term it is essential to strengthen the attractiveness of this particular career path, and a specific internal training program for young talent will soon be set up.

In addition to fundamental research, Idiap pursues important activities in the field of technology transfer,

"We are convinced of the relevance of our activities within the AI community"

Olivier Dumas
President of the Foundation Council, Idiap

and many start-ups have flourished thanks to this. And an additional path is opening up: the application of Idiap's technological expertise beyond the doors of the Institute. Following the creation of the Swiss Center for Biometrics in 2014, Idiap is now collaborating with the Centre de recherches énergetiques et municipales (CREM)—a non-profit body active in the field of energy in the urban domain—thus combining the knowhow of the two bodies. Further, Idiap now offers CREM its services and skills in a number of areas: human resources, communication, project management, and accounting and financial services.

More than ever before we are convinced of the relevance of our activities within the AI community. While some see the onset of an enslavement of humankind, we are convinced that artificial intelligence poses no specific threat. The only real danger lies in its misuse, and we should not forget that each time society acquires a more advanced technology science is accused of a multitude of evils. For our part, we dare to believe that humanity will, as before, be wise enough to not fall into unfortunate ways.

What makes Idiap so successful?

A t the invitation of a large European industrial group, I recently gave a presentation on the Idiap Research Institute, in the process learning that the group's extensive needs regarding artificial intelligence (AI) matched perfectly our own R&D offering.

This is true because today's Idiap is the result of building on solid foundations. Each stone that makes up those foundations is the result of the work of our researchers, research assistants, and doctoral students, and these foundation stones have been patiently laid for more than 25 years. The Institute ceaselessly offers newcomers the opportunity to develop their own activities, building them upon these foundations. Thanks to its long experience and the effective exploitation of knowledge acquired—in terms both of basic research and of the transfer of technology to industry—Idiap has become a reference when it comes to placing artificial intelligence (AI) at the service of society. The origin of a large number of software packages and databases used all over the world, Idiap demonstrates an exemplary command of the resources at its disposal.

At the level of fundamental research into AI, recent breakthroughs have proved no match for the groundbreaking findings of the 80s and 90s. Yet new opportunities and the renewal of interest in the domain lie in the exploitation of advances in ever-expanding storage and calculation capacities. And we can add to this the virtually unlimited amount of available data—Big Data. So now, what was conceived of more than twenty years ago is gently becoming reality, at the same time raising new challenges, both scientific and ethical.

One of the Institute's strengths is the coherence and independence of its research groups. Currently nine in number, they cover the majority of areas key to AI, which makes them particularly productive and competitive when they have the chance to work together on joint projects.

"Idiap, a one-stop shop for artificial intelligence"

Hervé Bourlard
Director, Idiap; Professor, EPFL

5

Thus, choosing to create a new group is always a risk. Idiap does not wish to simply follow fashion regarding the direction its research takes. To maintain its level of excellence, it must diversify in and around its chosen areas of expertise, while simultaneously anticipating future needs.

The creation in 2015 of the Uncertainty Quantification and Optimal Design group is a case in point. While the current trend exploits Big Data in particular, this research group is interested in solving problems with only a small amount of data available—"small data"—where one must develop a model from only a small quantity of indications. The knowledge developed in this area should be applicable in a large number of situations, including in the optimal placement of sensors for the evaluation of environmental risk, for example.

And this is how Idiap—based on its in-depth knowledge, the dynamism of each of its research groups, and its R&D team, which ensures the quality of its software packages and their transfer to industry—has created a unique offering in the field of AI.

organization

06 — 17

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.

I n 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 tools

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, 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 in France, England, and Germany. Across the Atlantic, Idiap and the International Computer Science Institute (ICSI) in Berkley, California, share a privileged partnership.

7

Structure

• Non-profit foundation

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

• Specializes in artificial intelligence, management of multimedia informations and man–machine, multimodal interactions

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

Financing

5% provisions

45% public funds

50% research projects
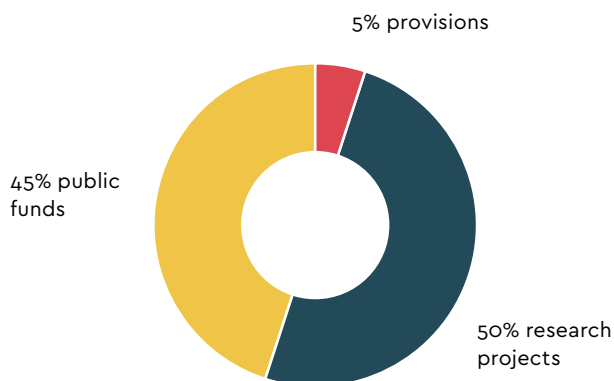
The three missions of Idiap

The Idiap Research Institute is recognized by the Swiss confederation as part of the strategic domain of the Swiss federal Institutes of technology and is supported under the Federal Act on the Promotion of Research and Innovation (RIPA) (Article 15). The Institute's vision is to promote quality of life through scientific progress made in relation to artificial intelligence (AI). Idiap's activities are oriented in accordance with three complementary missions:
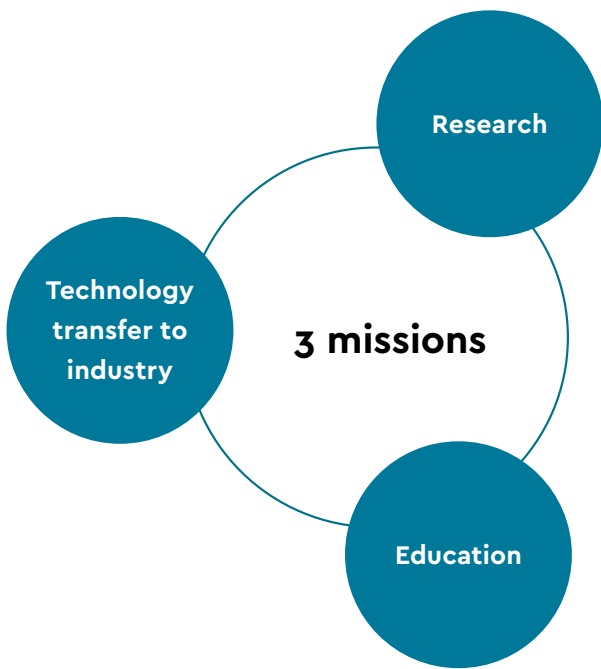
Research

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 world-wide. 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 around 50 percent of the Institute's financing.

Education

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.

Technology transfer to industry

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 around 10 percent of Idiap's budget. Thanks to the incubator—The Ark at IdeArk, Idiap contributes to the creation of numerous successful start-ups.

Research

Technology transfer to industry

3 missions

Education

Idiap in figures 2017

Human resources

4 professors

2 senior scientists (MER)

6 permanent senior researchers

7 research associates

18 postdocs

33 research assistants

11 development and research engineers

7 system engineers

7 administrative staff

13 trainees (average/year)

144 **people in total**

98 **full-time equivalents**

In total, 25 nationalities are represented at Idiap

- Switzerland
- France
- Rest of Europe
- India
- America
- Rest of Asia
- Stateless, 4 robots

Scientific activities

- Participation in **sixty-five research programs**

- Project management in more than **forty-five 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 company IdeArk

- **More than fifty posts** in IdeArk start-ups

Submission and financing of research projects

In 2017 Idiap submitted 52 projects (22 to Swiss funding agencies, 19 to European agencies, and 10 industrial projects and 1 international project). At the same time the Institute managed 65 active projects.

Publications

In 2017, Idiap researchers participated in 102 peer-reviewed publications.

Theses completed — 5

Conference articles — 64

Book chapters — 3

Scientific articles — 30

52 projects submitted in 2017

65 active projects in 2017

5 research themes
10 application domains

Perceptual and cognitive systems

Automatic speech recognition, computer vision, handwriting recognition, multimedia document processing, robotics, machine translation.

Social/human behavior

Web and mobile communication, social interaction, social signal processing.

Information interfaces and presentation

Multimedia information systems, personalization and contextualization.

Cyber-security

Person authentication; voice, face, and vein recognition; detection of criminal activity.

Machine learning

Statistical modeling, neural networks, mathematical models.

Human–human, human–machine & human-robot interaction

Voice-controlled devices, voice–voice translation systems, navigation systems, gesture synthesis, learning from demonstration.

Exploitation of rich multimedia archives

Semantic indexing, object recognition, audio–video content filtering, YouTube document analysis, analysis of cultural heritage media.

Collaborative and creative systems

Smart meeting rooms, video conferencing, multimedia indexing and access, cross-lingual collaboration.

Management of mobile systems

Signal processing for mobile platforms, mobile social networks, modeling of human activity to anticipate needs.

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, bio-imaging, microscopy.

Security and risk management

Access control (for physical or virtual spaces), speaker identification, face identification, video surveillance, natural-risk modeling, industrial engineering, safety of public and private spaces.

Cities of the future ("smart cities")

Ecology, environmental management, pollution reduction, traffic reduction, better use of road networks, noise reduction.

Entertainment and games

Multi-lingual gaming, remote-family games ("togetherness").

Home automation (domotics)

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

Organization chart

Foundation Council

Scientific Advisory Board

Director
Hervé Bourlard

Deputy Director
François Foglia

Scientific College
F. Fleuret, D. Gatica-Perez, J.-M. Odobez

Admin. Committee

Data and Research Ethics Committee

Research Committee

11

Director, Finance & Human Resources
Ed Gregg

System & Infrastructure
Frank Formaz

Speech and Audio Processing
Hervé Bourlard
Phil Garner
Mathew Magimai Doss
Petr Motlicek

Robot Learning and Interaction
Sylvain Calinon

Legal Adviser
Marie-Constance Kaiflin

Research & Development
Olivier Bornet

Social Computing
Daniel Gatica-Perez

Uncertainty Quantification and Optimal Design
David Ginsbourger

Program Managers
Léonore Miauton

Technology Transfer
Florent Monay

Machine Learning
François Fleuret

Natural Language Understanding
James Henderson

Communication, Public Relations & Marketing
François Foglia

Perception and Activity Understanding
Jean-Marc Odobez

Computational Bio-Imaging
Michael Liebling

**Biometrics Security and Privacy**
Sébastien Marcel

Research groups

Nine groups, each led by one or more top-level scientists, divide the Institute's research activities between them. The group responsible for speech- and audio processing is led by the Institute's director – Hervé Bourlard – and three permanent researchers. The remaining groups are all headed by permanent researchers, including one tenured professor and two senior scientists.

Speech & Audio Processing
Prof. Hervé Bourlard (second from right), Dr. Phil Garner, Dr. Petr Motlicek, Dr. Mathew Magimai-Doss / H-indexes*: PG—23, PM—16, HB—59, MMD—25
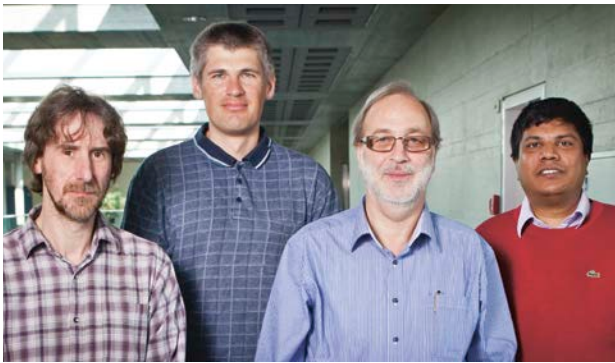
Speech processing has been one of the mainstays of Idiap's research portfolio for many years. Today it is still the largest group in the Institute, and Idiap continues to be recognized as a leader in the field. The expertise of the group encompasses statistical automatic speech recognition (based on 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
Prof. Daniel Gatica-Perez
h-index: 55

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: 29

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.

* **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.

12

Perception & Activity Understanding
Dr. Jean-Marc Odobez
h-index: 40

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.

Robot Learning & Interaction
Dr. Sylvain Calinon
h-index: 35

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.

Uncertainty Quantification and Optimal Design
Prof. David Ginsbourger
h-index: 22

The Uncertainty Quantification and Optimal Design group focuses on quantifying and reducing uncertainties in the context of hi-fidelity models, with a main expertise in Gaussian process methods and the sequential design of computer experiments for optimization, inversion, and related problems. Application domains notably include energy and geosciences, with current collaborations ranging from safety engineering to hydrology and climate sciences.

Computational Bio-Imaging Group
Prof. Michael Liebling
h-index: 21

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.

Biometrics Security and Privacy
Dr. Sébastien Marcel
h-index: 43

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 Understanding
Dr. James Henderson
h-index: 22

The Idiap Natural Language Understanding (NLU) group was created in September 2017 under the direction of James Henderson, in part as a continuation of the previous Natural Language Processing group, which was lead by Andrei Popescu-Belis. The NLU group continues work on how semantic and discourse processing of text and dialog can improve statistical machine translation and information indexing, with a recent focus on neural machine translation and attention-based deep learning models. This fits well with the NLU group's new research direction of neural-network-structured prediction and representation learning for modeling the syntax and semantics of text and speech, including modeling abstraction (textual entailment) and summarization.

Foundation Council

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.

Mr. Olivier Dumas, President
Independent manager and business consultant

Mr. Jordi Montserrat
Regional Manager Venturelab

Ms. Anne-Laure Couchepin Vouilloz, Vice-President
President of the Town of Martigny

Prof. Stéphane Marchand-Maillet
Associate professor at the Department of Computer Science of the University of Geneva

14

Mr. Marc-André Berclaz
Valais–Wallis Cluster

Mr. Dominique Perruchoud
President of the Board of Cimark SA

Mr. Stefan Bumann
Head of the Service des Hautes Ecoles (SHE)

Prof. Pierre Vandergheynst
Professor and Vice-President for Education at EPFL

Mr. Patrick Furrer
Scientific collaborator at swissuniversities

Mr. David Watrin
Head of the product unit Security & Intelligence at Swisscom

Advisory Board

The Advisory Board is composed 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 support is frequently sought and proves to be valuable when making decisions regarding research, training, and technology transfer.

Dr. Alex Acero

Senior Director at Apple,
Cupertino, CA, USA

Prof. Klaus-Robert Müller

Professor of Computer Science,
TU Berlin; Director, Bernstein
Focus on Neurotechnology,
Berlin, DE

15

Prof. Dr. Vanessa Evers

Full Professor of Human Media
Interaction, University of
Twente, NL

Dr. Prem Natarajan

Director of the Information
Sciences Institute, affiliated with the Viterbi School
of the University of Southern
California, USA

Prof. Anil K. Jain

Distinguished Professor,
Department of Computer
Science & Engineering, Michigan
State University, USA

Prof. Bernt Schiele

Max Planck Institute, Director,
MPI Informatics; Professor
at Saarland University,
Saarbrücken, DE

Prof. Johanna Moore

University of Edinburgh, Head
of the School of Informatics;
Director of the Human
Communication Research
Center, UK

Dr. Luciana Vaccaro

Rector of the HES-SO, University
of Applied Sciences and Arts,
Western Switzerland, CH

Employees

Scientists
First name, last name, position, country of origin,
year of arrival

Afsaneh Asaei, postdoc, Iran, 2008

Sushil Bhattacharjee, research associate, Switzerland, 2016

Hervé Bourlard, director, Belgium/Switzerland, 1996

Sylvain Calinon, researcher, Switzerland, 2014

Yuanzhouhan Cao, postdoc, China, 2017

Gülcan Can, research assistant, Turkey, 2013

Olivier Canévet, research assistant, France, 2012

Tatjana Chavdarova, research assistant, Macedonia, 2014

Tiago De Freitas Pereira, research assistant, Brazil, 2014

Subhadeep Dey, research assistant, India, 2014

Pranay Dighe, research assistant, India, 2013

Pavankumar Dubagunta, research assistant, India, 2017

François Fleuret, senior researcher, France, 2007

Philip Garner, senior researcher, United Kingdom, 2007

Daniel Gatica-Perez, senior researcher, Mexico, 2002

David Geissbuhler, postdoc, Switzerland, 2017

Anjith George, postdoc, India, 2017

David Ginsbourger, senior researcher, France, 2015

Weipeng He, research assistant, China, 2016

James Henderson, senior researcher, United States of Amercia, 2017

Guillaume Heusch, research associate, Switzerland, 2015

Rui Hu, postdoc, China, 2013

David Imseng, postdoc, Switzerland, 2009

Christian Jaques, research assistant, Switzerland, 2016

Noémie Jaquier, research assistant, Switzerland, 2016

Cijo Jose, research assistant, India, 2013

Angelos Katharopoulos, research assistant, Greece, 2017

Ina Kodrasi, postdoc, Albania, 2017

Alain Komaty, postdoc, France, 2016

Pavel Korshunov, postdoc, Estonia, 2015

Vedrana Krivokuca, postdoc, New Zealand, 2017

Thibaut Kulak, research assistant, France, 2017

Andras Kupcsik, postdoc, Hungary, 2017

Florian Labhart, research assistant, Switzerland, 2017

Nam Le, research assistant, Vietnam, 2015

Yann Lepoittevin, postdoc, France, 2016

Michael Liebling, senior researcher, Switzerland, 2014

Gang Liu, postdoc, China, 2017

Srikanth Madikeri, postdoc, India, 2013

Mathew Magimai Doss, researcher, India, 2007

Sébastien Marcel, senior researcher, France/Switzerland, 2000

Olivia Mariani, research assistant, Switzerland, 2016

Angel Martinez Gonzalez, research assistant, Mexico, 2016

Amir Mohammadi, research assistant, Iran, 2016

Petr Motlicek, researcher, Czech Republic, 2005

Hannah Muckenhirn, research assistant, France, 2015

Skanda Muralidhar, research assistant, India, 2014

James Newling, research assistant, United Kingdom, 2013

Laurent Nguyen, postdoc, Switzerland, 2011

Olegs Nikisins, postdoc, Latvia, 2017

Jean-Marc Odobez, senior researcher, France/Switzerland, 2001

Nikolaos Pappas, research assistant, Greece, 2012

Trung Phan, research assistant, Vietnam, 2015

Emmanuel Pignat, research assistant, Switzerland, 2016

Xiao Pu, research assistant, China, 2014

André Rabello Dos Anjos, research associate, Brazil, 2010

Dhananjay Ram, research assistant, India, 2014

Ramya Rasipuram, postdoc, India, 2010

Marzieh Razavi, research assistant, Iran, 2013

Bastian Schnell, research assistant, Germany, 2017

Jilt Sebastian, research assistant, India, 2017

Adrian Shajkofci, research assistant, Switzerland, 2016

Rémy Siegfried, research assistant, Switzerland, 2017

Suraj Srinivas, research assistant, India, 2017

Ajay Srinivasamurthy, postdoc, India, 2017

Ajay Tanwani, research assistant, Pakistan, 2015

Sibo Tong, research assistant, China, 2016

Sandrine Tornay, research assistant, Switzerland, 2016

Michael Villamizar, postdoc, Spain, 2017

Bogdan Vlasenko, postdoc, Germany, 2017

Yu Yu, research assistant, China, 2015

Research & Development Engineers
Philip Abbet, senior development engineer, Switzerland, 2006

Olivier Bornet, head of development team, Switzerland, 2004

Milos Cernak, senior development engineer, Slovakia, 2011

Guillaume Clivaz, development engineer, Switzerland, 2017

Samuel Gaist, development engineer, Switzerland, 2013

Mélanie Huck, development engineer, Switzerland, 2017

Salim Kayal, development engineer, Switzerland, 2011

Christine Marcel, development engineer, France/Switzerland, 2007

Florent Monay, tech transfer officer, Switzerland, 2008

Alexandre Nanchen, senior development engineer, Switzerland, 2008

Flavio Tarsetti, senior development engineer, Switzerland, 2008

16

Administrative staff

First name, last name, position, country of origin,
year of arrival

Elisa Bovio, junior program manager and financial assistant,
Switzerland 2015

Martina Fellay, program manager, Austria, 2012

François Foglia, deputy director, Switzerland, 2006

Edward-Lee Gregg, chief financial officer, United States of
America, 2004

Marie-Constance Kaiflin, legal adviser, Switzerland, 2017

Léonore Miauton, head of program management,
Switzerland, 2012

Sylvie Millius, secretary, Switzerland, 1996

Nadine Rousseau, secretary, Belgium/Switzerland, 1998

System engineers

Bastien Crettol, system administrator, Switzerland, 2005

Norbert Crettol, senior system administrator, Switzerland, 2002

Cédric Dufour, system engineer, Switzerland, 2007

Frank Formaz, system manager, Switzerland, 1998

Louis-Marie Plumel, senior system administrator, France, 2011

Vincent Spano, webmaster, Switzerland, 2004

Laurent Tomas, senior system administrator, Switzerland, 2017

Interns

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

**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 pro-
jects in which Idiap participates.**

Dylan Bourgeois, France, EPFL

Jaden Diefenbaugh, USA, Oregon Univ, Corvallis, USA

Selen Kabil, Turkey, Middle East Technical Univ, Ankara, Turkey

Wei Ma, India, Nanjing Univ, China

Luis Emmanuel Medina Rios, Mexico, EPFL

Zohreh Mostaani, Iran, Ozgegin Univ, Istanbul, Turkey

Amrutha Prasad, India, Karnataka, India

Alain Rossier, Switzerland, EPFL

Florian Standaert, Switzerland, EPFL

finances

18 — 21

Profit and loss statement (CHF)

INCOME	2016	2017	%
City of Martigny	700 000	**700 000**	6.3
Canton of Valais	1 720 000	**2 000 000**	17.9
Swiss Confederation	2 423 700	**2 418 500**	21.6
Loterie Romande	87 000	**90 000**	0.8
Swiss National Science Foundation projects	1 537 844	**1 268 033**	11.3
Hasler Foundation projects	175 781	**224 838**	2
European Commission projects	1 404 423	**1 760 239**	15.7
The Ark projects	245 380	**215 650**	1.9
CTI projects	638 467	**1 007 604**	9
US projects	—	**296 733**	2.6
EPFL contribution	72 000	**72 000**	0.6
Industrial financing	414 669	**147 015**	1.3
Other income/extraordinary income	738 592	**490 277**	4.4
Provisions	555 000	**510 000**	4.6
TOTAL INCOME	**10 712 856**	**11 200 889**	**100**

19

EXPENSES	2016	2017	%
Personnel expenses	7 948 334	**8 588 346**	76.7
Education and travel	427 372	**371 307**	3.4
Third-party expenses	102 292	**229 278**	2.1
Computer equipment and maintenance	253 410	**355 319**	3.2
Administrative costs	166 113	**158 719**	1.5
Promotion and communication	45 769	**18 147**	0.2
Rent	709 229	**671 754**	6
Depreciation	363 090	**482 636**	4.3
Exceptional expenses	—	**—**	—
Variation provision	685 000	**290 896**	2.6
TOTAL EXPENSES	**10 700 609**	**11 166 402**	**100**
OPERATING PROFIT/LOSS	**12 247**	**34 487**	**—**

Accounting analysis 2017

For the first time in its history, Idiap ends a fiscal year with revenues in excess of 11 million Swiss francs.

Public subsidies, proportionally always less than 50 percent of all revenues, are essential for the proper functioning of the Institute and for its development. An increase of the cantonal share of such subsidies, championed by the Service des Hautes Ecoles de l'Etat du Valais, will allow, in 2018, the creation of two new research groups that should generate further new revenues.

The growing budget for projects is funded notably from a new source of financing. Three new projects are now being funded by the US Department of Defense (DARPA) or US National Intelligence (IARPA).

Personnel costs (salaries plus social charges) make up approximately three-quarters of total expenses. For each franc of cantonal subsidies, leverage is currently greater than 5.5.

Idiap closes its 2017 accounts with a profit of CHF 34,487.

Federal, cantonal, and municipal subsidies

(In thousands of Swiss francs)

| YEAR	2015	2016	2017	2018*
Confederation	2 467	2 424	2 418	2 406
Canton	1 720	1 720	2 000	2 000
Municipality	700	700	700	700

*Budget

20

Distribution of funding sources

- City of Martigny
- Canton of Valais
- Swiss Confederation
- Loterie Romande
- Swiss National Science Foundation projects
- Hasler Foundation projects
- Eur Commission projects
- The Ark projects
- CTI projects
- US Projects
- EPFL contribution
- Industrial financing
- Other/extraordinary income
- Provisions
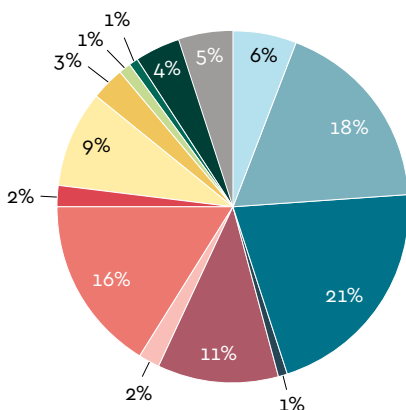
Distribution of Costs

- Personnel expenses
- Education and travel
- Third-party expenses
- Computer equipment and maintenance
- Administrative costs
- Promotion and communication (0%)
- Rent
- Depreciation
- Exceptional expenses (0%)
- Variation provisions
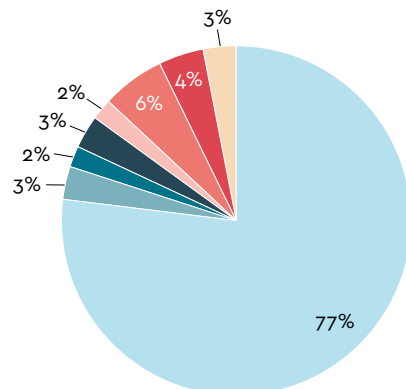
Balance sheet (CHF)

ASSETS	31.12.2016	31.12.2017
Cash	5 313 117.83	5 306 132.52
Accounts receivable	130 394.85	61 886.00
Accrued income and other	434 238.84	769 938.24
TOTAL CURRENT ASSETS	**5 877 751.52**	**6 137 956.76**
Equipment	590 125.14	458 112.24
Financial assets	10 000.00	10 000.00
TOTAL NON-CURRENT ASSETS	**600 125.14**	**468 112.24**
TOTAL ASSETS	**6 477 876.66**	**6 606 069.00**

| LIABILITIES	31.12.2016	31.12.2017
Accounts payable	213 106.26	472 526.46
Accrued expenses	3 747 835.83	2 517 150.17
Provisions	1 085 000.00	1 450 000.00
TOTAL FOREIGN FUNDS	**5 045 942.09**	**4 439 676.63**
Share capital	40 000.00	40 000.00
Research funds reserve	0.00	700 000.00
Special reserve	1 200 000.00	1 200 000.00
Retained earnings	179 687.15	191 934.57
Net income	12 247.42	34 457.80
TOTAL OWN FUNDS	**1 431 934.57**	**2 166 392.37**
TOTAL LIABILITIES	**6 477 876.66**	**6 606 069.00**
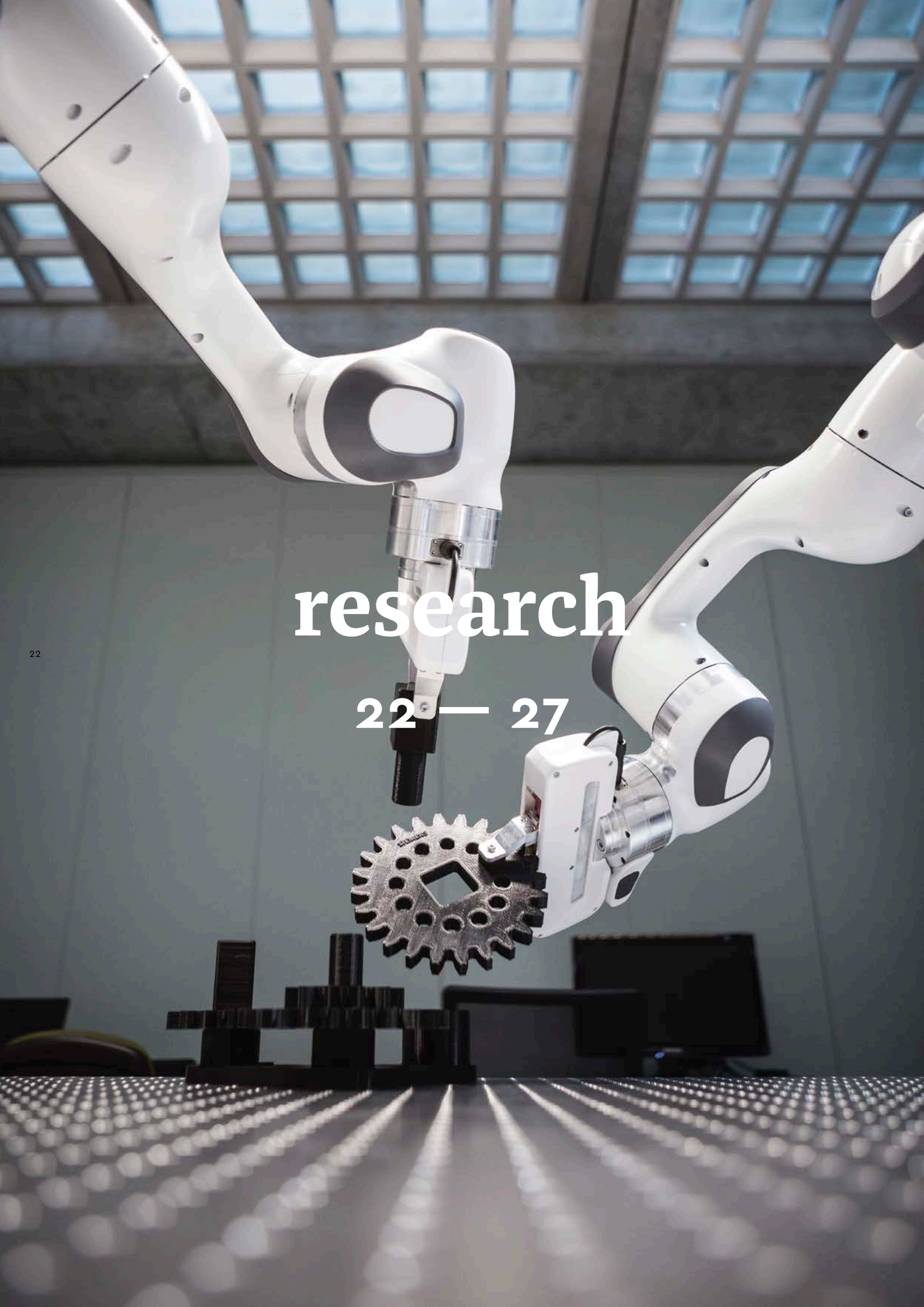
research

22 — 27

MEMMO—Memory of motion to enable rapid adaptation in robots

For each of us, navigating around obstacles that we find in our path is a natural activity. But for an exoskeleton used, for example, by a paraplegic person, or for a quadruped robot this is much more of a challenge. To maintain its balance and control it has to rapidly calculate its options. A response that is too slow in coming or a poor choice made inevitably results in a fall. The solution? Integrate, into the robot's programming, memories of similar experiences of maintaining balance. And this is precisely what Idiap is addressing in the European program MEMMO.
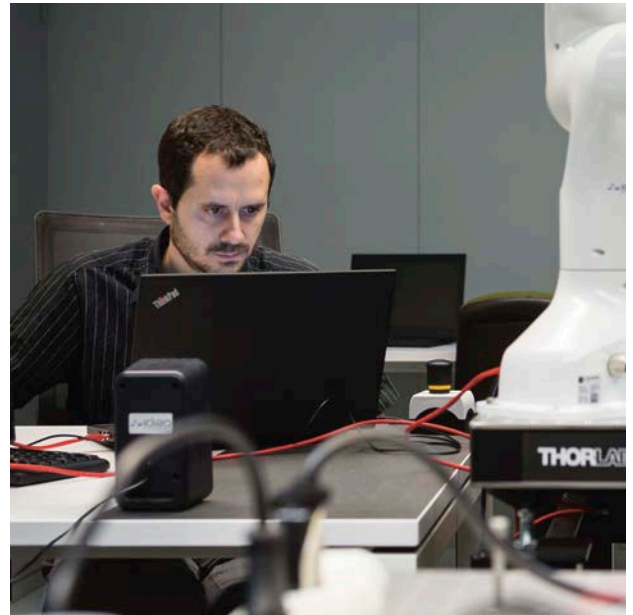
MEMMO (Memory of Motion) is a major European project with a budget of EUR 4 million. Lasting 48 months, it brings together a consortium of academic and industrial stakeholders, who will work on finding ways to "give" robots memory.

Under the coordination of an independent unit of the French National Center for Scientific Research (CNRS), Dr. Sylvain Calinon, senior researcher and head of Idiap's Robot Learning and Interaction group, will lead the Idiap teams. The other academic partners are the universities of Edinburgh and Oxford, as well as the Max Planck Institute, Tübingen.

But the consortium doesn't confine itself to university labs; it also involves industrial partners from across Europe, each with its know-how, but also its specific goals. PAL Robotics, from Barcelona, specializes in the creation of humanoid robots; Wandercraft, based in Paris, develops exoskeletons specially designed for paraplegic people. Airbus, too, has joined the project—the aircraft manufacturer needs

anthropomorphic robots to perform certain operations inside the cabins of its aircraft. And Britain's Costain Group wants to develop quadruped robots to inspect piping networks.

What are the project's challenges? It's difficult for a robot to learn when it only has limited experience of the real world. This learning is all the more vital for humanoid robots, which will certainly be required to perform varied tasks in a range of different environments. Fortunately, further experience can be gained via simulations, in which the robot—safe in a virtual environment—"imagines" incidents that might occur and tries to cope with them. Successful experiences are then analyzed and stored as a memory of motion.

So, thanks to the research carried out in the MEMMO project, these and other robots will be able to better handle both changes to their environments and unforeseen events.

Martigny, the heart of a European language-study network

TAPAS (Training Network on Automatic Processing of Pathological Speech) is an unusual project. Supported by a highly selective financing mechanism established by the European Union, the Idiap-led speech research program focuses on the international and cross-sectoral mobility of researchers from all walks of life. A major, ambitious challenge that has propelled IDIAP into the very heart of advanced "Made-in-Europe" science.

The idea for the project emerged from the long-standing collaboration between Dr. Mathew Magimai-Doss, senior researcher at Idiap and coordinator of TAPAS, and Dr. Heidi Christensen of Sheffield University. To accompany them on this significant adventure and to enable them to meet the project's various research objectives, twelve academic and industrial institutions and nine partner organizations have joined the consortium—with Idiap at the helm.

A project of exceptional size requires exceptional financing: TAPAS has attracted funding from the well-known Marie Skłodowska-Curie Actions (MSCA) grant program, which particularly encourages training networks in cutting-edge scientific fields, focusing on both geographical and cross-sectoral mobility. Something that Dr. Magimai-Doss is extremely proud of: "MSCA-type financing is highly sought after and involves fierce competition between the projects submitted."

Speech impairment and the need for a multidimensional approach

A number of pathologies can affect a person's ability to express themselves, including Alzheimer's, Parkinson's, and stroke. Those concerned often lead a life hallmarked by exclusion.

The TAPAS project will study this problematic from three main angles: detection of language impairment, evaluation and therapeutic tools, and rehabilitation. Dr. Magimai-Doss and the project partners have devised a multidisciplinary approach that ranges from the most fundamental research to the most practical of applications. The common goal of all of these elements is, as the coordinator of TAPAS puts it, "to understand how a pathology affects the production and the perception of spoken language."

Fifteen doctoral students, fifteen facets of speech processing

As part of the project, a consortium of clinicians, research groups, and industrial partners will engage fifteen PhD students across a range of European locations. In addition to their three-year course of scientific study, these students will receive theoretical and practical training, ranging from questions of ethics and legal, data-centered issues to entrepreneurship training. The scientific objective will be to study the different facets of speech impairment and the condition's impact on communication.

A first group of students will measure the intelligibility of people suffering from a pathology. As the traditional methods of evaluating human speech do not give satisfactory results in cases of disability, this will involve—for example—developing new intelligibility scores.

Another group will use the study of language, and especially of the underlying causes of problems with speech, to detect degenerative diseases and measure their evolution.

Finally, at the heart of another research theme we find the question of how best to help those affected by a pathology through speech technology. For example, the modeling of children's language will enable the creation of a therapeutic tool in the form of a game for children with speech disorders.

Idiap knows all the moves

An unprecedented new platform for capturing moving objects and validating research findings
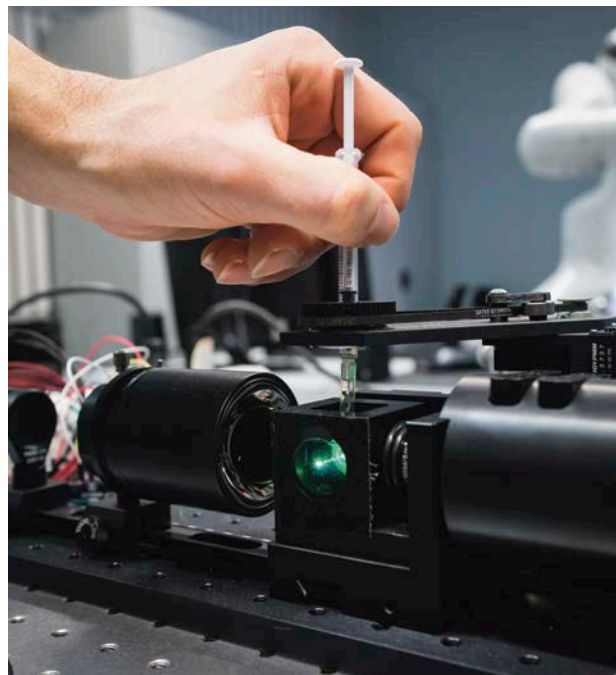
Why a multimodal data (MMD) platform?

"There's a vacuum with regard to the acquisition of reproducible data from moving objects, from moving sensors, or in changing lighting conditions. The idea for the platform came out of a discussion between researchers, with various groups at Idiap believing it was necessary," explains Dr. Michael Liebling, Head of the Computational Bio-Imaging Group. In collaboration with two other Idiap senior researchers, Sylvain Calinon and Sébastien Marcel, Liebling submits a grant request to the Swiss National Science Foundation (SNSF) late in 2015. That request is accepted the following year.

MMD complements existing research activities, including the high-performance BEAT platform, which processes biometric data.

What is the goal of this type of platform?

"When researchers use algorithms in their research, how can they be sure that the results are accurate? Will it be possible to reproduce these results on the other side of the world?" Liebling asks. Because, of course, one of the conditions necessary for the validation of research findings is that they are reproducible.

Idiap supports such validation efforts through the quality of its annotated databases, which allow experiments to be replicated with established algorithms and give researchers the ability to evaluate and compare the quality of new algorithms.

Building a microscopy system "made in Valais"

The Idiap MMD platform includes "an open-SPIM microscopy system dedicated to the acquisition of 3-D images, for which it was necessary to build custom-made parts." The "poly-mechanics" of Sion airbase not only made the parts, their suggestions also helped the project avoid certain pitfalls during the assembly phase.

Thanks to this new microscopy system, which allows the creation of comparison points, computational bio-imaging researchers can now validate their algorithms. The system can reconstruct, for example, images of a beating heart with unmatched precision. Liebling is looking forward to all the new opportunities this platform will offer his research, and that of his colleagues.

Not far away, Sylvain Calinon and his team have placed a 3-D image capture device in the hand of one of the two robotic arms offered by the platform. The goal is to reconstruct the arm's movement using the images captured at the hand. Will the algorithms used prove effective? MMD will answer that question, and many more besides.

Another of the platform's goals is to build and share databases in areas that are currently lacking such tools. So, with this new platform—born of a creative environment and the result of multiple collaborations—Idiap researchers once again find themselves at the forefront of movement.

Artificial intelligence in the research spotlight

Machine learning makes for a full house

Machine learning, the "field of study that gives computers the ability to learn without being explicitly programmed", is nothing new—American Arthur Samuel already laid down a provisional definition in 1962. More than half a century later, artificial intelligence (AI) has once more become a topic of interest in research circles, and a central issue for society. François Fleuret, Head of the Machine Learning group at Idiap and a lecturer at EPFL specializing in neural networks, has organized the Swiss Machine Learning Day every year since 2012. The event brings Swiss researchers together to explore the theme and attracts, year after year, an ever larger audience. Last November, a victim of its own success, one of the largest halls of the SwissTech Convention Center in Lausanne was required to welcome the many guests.

Torch, the Swiss army knife of machine learning—developed at Idiap

In order to solve complex issues involving large amounts of data, machine learning algorithms require a very flexible support. In 2001, Idiap answered this need, developing Torch—a network of "deep learning" libraries, ideal for "simulating neural networks". Almost two decades later, Torch is a fully operational tool, used in particular by Facebook. And Facebook has also developed a direct descendant of Torch—Pytorch—adopted among others by Uber, and many researchers around the world.

Idiap and HES-SO Valais/Wallis—working hand in hand

Open to the public and the scientific community, researchers from the two Valais institutes organize joint workshops on artificial intelligence, offering—with the aim of stimulating collaboration and exchange—two sessions each year. These free-of-charge initiatives are intended as universal platforms for the dissemination of knowledge. And the company Klewel, an Idiap spin-off, makes the events available to a broader public online, with the option of watching them speaker by speaker.

The theme of the first workshop of the year was issues related to the reproducibility of scientific research. Initiated by Dr. Sebastien Marcel, Head of the research group Biometrics Security and Privacy, this hot topic was discussed and debated throughout the day. A distinguished guest, EPFL's Vice-President for Education Prof. Pierre Vandergheynst, presented the EPFL position and the Federal Institute of Technology's approaches to meeting this challenge.

The second workshop was organized by Professor David Ginsbourger, Head of Uncertainty Quantification and Optimal Design at Idiap. This time the focus was on "Small Data" and the issue of how research sometimes has to deal with constraints on data availability, as is the case with research in the field of rare diseases, or regarding ecological, climate, or financial disasters and crises. Even Idiap's robotics researchers are confronted with the obstacles created by Small Data. An unusual subject, and one that touches on models of pure mathematics.

300 people at Martigny for "AI in all its forms"

Representatives of the private and public sectors of the regions crossed by the Rhône river met in Martigny at the Vatel Hotel on September 15, 2017 for the 17th edition of the Rhône Forum. The goal: to promote the exchange of ideas and to support project opportunities involving economic actors from Switzerland and neighboring France. The event, sponsored by Business Valais and the city of Martigny, focused on

the technological innovations brought about by AI. Director of Idiap Professor Hervé Bourlard shared his views on the subject during his presentation, "Artificial Intelligence—Between Reality and Fiction, between Research and Propaganda". Senior Idiap researchers Sébastien Marcel and Jean-Marc Odobez also took part in round tables dedicated to cybersecurity and to robotics. This event made possible useful exchanges between science, the political and economic authorities, and the general public.

network

28 — 33

International Create Challenge: 6th edition

This three-week, total-immersion program enables young entrepreneurs to transform their ideas into commercial prototypes. Running from 30 August to 20 September 2017 at the Idiap Research Institute in Martigny, it brought together participants from a number of countries, including Brazil, Ukraine, China, and Argentina. The ICC jury—made up of a dozen internationally renowned figures from the worlds of research, venture capital, and innovation—awarded prizes to the following teams. Each win a CHF 5,000 cash prize:

- **Forensic Voice Recognition System** (Srikanth Madikeri, Idiap) is developing a tool to facilitate the identification of individuals based on vocal biometrics. In addition to the aforementioned prize, the project qualifies for free access to the incubator The Ark.

- Robobay (Yaroslav Pishta, Ukraine) is an Internet platform that allows its clients to deploy their own trading bots to interface with the market for crypto-currency.

- TOP DOWN (Renato Rodrigues, Brazil) is developing a payment and use-monitoring technology that facilitates use of public transport.

"The Challenge is made possible by the commitment of those experts that assist its participants"

The ICC, a proven technology transfer tool

The International Create Challenge (ICC) is at the origin of companies such as Recapp, Biowatch, EyeWare, and Anemomind. With regard to the 2017 edition, negotiations are under way with several teams concerning their integration into the incubator The Ark, situated on the IdeArk site in Martigny.

François Foglia, deputy director of Idiap and creator of the ICC, highlights the important role played by the experts who have been involved in this adventure for the past six years. On each day of the Challenge, one of these experts graciously shares his or her knowledge or coaches the teams of budding entrepreneurs.

Christophe Saam (P&TS) is present every year to explain the rules of intellectual property. Gaspard Couchepin, barrister, notary, and Doctor of Law, explains the various types of company. Laurent Bischof and Guillaume Dubret (Polytech Ventures) talk on venture capital practices. Nadine Reichenthal, another expert and a lecturer in Strategy at the University of Lausanne, talks about the "business canvas" and the value proposition. Jordi Montserrat (Venture Kick) shares his experience of launching start-ups. Jung Park (IMD), Joel Rossier and Daniel Rüfle (The Ark), Jasmine Abarca-Golay (InnoPark), Patrick Biro (Venture Kick), and Stéphane Marchand-Maillet (University of Geneva) also support the participants each year or contribute to the work of the jury.

Reaching out to the wider public—a vocation for Idiap

The *"Carnotzet Scientifique"*—research in good company

Free and open to all, the *"Carnotzet Scientifique Valaisan"* places the wider public and invited experts on an equal, friendly footing. A panel of experts met at the bookstore La Liseuse in Sion on May 10, 2017 to discuss the theme "Serious Game: Using a Playful Approach to Solve Very Serious Problems". This *Carnotzet Scientifique* was the perfect opportunity for researchers to interact with members of the general public in a relaxed, informal atmosphere. Professor Daniel Gatica-Perez—Director of Idiap's Social Computing group—also gave a presentation, on the smartphone as an often-underestimated tool of civic participation, an approach based on observations made during his research at Idiap, at the crossroads of computers, social media, machine learning, and the social sciences.

Pepper, a YouTube star from Valais

Pepper, one of Idiap's robots, was invited to speak at the TEDx Martigny on September 8, 2017. Before a packed house, he explained—not without humor—the reason for his presence: "The organizers couldn't find a human fool enough to talk about artificial intelligence. So, as is the case for so many things that the enlightened biped doesn't like to do, it's me who gets given the job." One of the most memorable interventions and one that continues to gather Internet followers: Pepper's speech, available on YouTube, has already racked up more than 2,400 views.

At the beginning of November, the disarming robot charmed at the State of Valais Cultural Awards ceremony. Pepper stole the limelight by interviewing Valais actress Mali Van Valenberg, who unfortunately could not be there in person but was able to express her gratitude via Pepper's control screen. Adopted without hesitation by the other laureates present, Pepper even posed for the group photo. A great success for Idiap's researchers, and an intriguing discovery for the public.

Come along and visit Idiap!

Idiap's doors are open to all comers every second Wednesday of the month from 5:00 to 6:00 pm (or at other times on request). François Foglia, deputy director of the Institute, presents the activities of the research groups and gladly answers the questions of the most curious visitors.

Activities
for young people

The future of research meets
Future en tout genre

Since 2012, the Institute has been welcoming children for one morning each year as part of the program "Future en tout genre". Last November, nineteen children aged 11 or 12 came to Idiap to explore different elements of the jobs of researchers, developers, network managers, or project managers. And as well as participating in the workshops organized for them, the children got to meet Nao and Pepper, Idiap's robots. The initiative, launched by the Valais Secretariat for Equality and the Family, offers children the opportunity to consider their futures with any potential gender barrier removed. By broadening the field of opportunity open to them, Futur en tout genre helps children to understand, and respond to, their own inner drives. It will be a few years, however, until we know if any of 2017's visitors have fallen for the sciences.

Hérisson sous gazon
in Charrat

This open-air festival for children from 3 to 14 years of age takes place each year in Charrat, attracting over 7,000 visitors over a June weekend. Ever present, Idiap never misses the opportunity to put on a fun workshop where young and old alike can discover and interact with robots. A fine opportunity to showcase the Institute's research into signal recognition, whether those signals be auditory, visual, social, or behavioral. And the youngest in the audience are always particularly happy to talk with Nao, the smallest of Idiap's robots.

Idiap spin-offs—influence without frontiers

B iowatch (www.biowatch.ch/), Recapp (www. recapp.ch/), Eyeware (eyeware.tech/), and Koemei. Four start-ups, four success stories, one thing in common: Idiap labs. Spotlight on 2017, a year synonymous with success, for these spin-offs and for the Institute as a whole.

All doors are opening for Biowatch

It all began in 2014. Matthias Vanoni, working on his doctoral thesis, takes on a crazy challenge: to create a system that opens any door, even virtual, thanks to a sensor placed on the veins of the wrist (a technique developed in the laboratories at Martigny). After a first phase, in which the project receives support via the International Create Challenge program (http://www.createchallenge.org/), Biowatch tastes real success in 2017: it wins first prize at the Swiss Fintech Convention in Geneva, and—on the fringes of the renowned World Mobile Congress in Barcelona—Visa's Everywhere Initiative, which aims to identify the next big thing in payments. The spin-off has also won several other prestigious awards, including from Thales, LVMH, Accenture, and Credit Suisse. Another piece of news that's proving appealing to potential investors: in 2017 Biowatch was awarded the CTI Start-up Label, proof of its solidity and maturity, despite its young age.

Valais at the Swiss ICT awards ceremony 2017

In Lucerne last November, two Idiap spin-offs distinguished themselves at the Swiss ICT awards ceremony, an annual event that recognizes the most innovative ICT products and services in Switzerland. Recapp, which specializes in latest-generation audio recording and transcription, picked up the Newcomer Award and climbed to second place in the organizers' top 100. The spin-off is no stranger to the Swiss market,

collaborating with the Canton of Valais' Parliament (MediaParl project) and Swisscom, among others. Finally, Eyeware, an innovative solution enabling visual interaction between computers and people with disabilities (ICC'2015 award), received a nomination.

A buy that will make itself heard

The Swiss company Crealogix, specialist in FinTech solutions for banks, has chosen to use the technology developed by Idiap spin-off Koemei. Created in 2010, Koemei was built on the Institute's research into voice recognition, text analysis, and natural language, and sorts content automatically. The solution has been used by several universities in the context of higher education offerings.

32

Idiap alumni, a strong bridge between research and industry

To go from labs to industry takes only one step. And several at Idiap have not hesitated to take that step, and have done so with great success. Milos Cernak, for example, is now Technical Director at Logitech, and Alexandros Lazaridis is Director of Speech Processing in the Artificial Intelligence and Machine Learning division at Swisscom.

Idiap's dynamic ecosystem provides a fertile ground for collaborations between researchers. Beyond a simple "atmosphere", there is a genuine entrepreneurial mindset omnipresent throughout the corridors of the Institute—a mindset that also reaches out from those corridors, as is the case for Hugues Salamin and Yann Rodriguez, both currently at KeyLemon.

Cooperation and proximity—two vital factors in Idiap's success

At the time—from 2007 to 2009—a PhD student in Social Signal Processing, Salamin meets Rodriguez, one of the founders of KeyLemon. Rodriguez has completed his PhD and is already working with developers on the project. The two researchers hit it off immediately, and no one doubts that this first meeting will influence their respective futures, and that of the spin-off.

Salamin leaves for a while for Scotland, but returns to the Institute in 2014 in the role of developer, taking charge of the technology transfer office—a post that gives him solid experience of developing projects with industry and enables him to enrich his professional network. That profile appeals to KeyLemon, then growing rapidly, and the spin-off brings Salamin on board in 2017.

Now a product manager, the *Valaisan* extols the virtues of the proximity created and nurtured at Idiap: "Everyone knows each other, and that significantly facilitates collaboration because the expectations from all sides are so clear. We needed an outside evaluation of our mobile phone solution, and it was an Idiap postdoc who carried that out. It allowed us to move faster."

And Salamin also underlines the significance of him having found a post in Valais in his field of expertise—proof that the Institute is an attractive recruitment pool for companies in the canton.

33

faces

34 — 39

JAMES HENDERSON

From MIT to Martigny—the odyssey of a globe-trotting researcher with a fascination for language

American researcher James Henderson joined Idiap with brio last autumn. Leading the research group Natural Language Understanding, his focus includes the semantic processing of texts and speech with the aim of improving automatic translation. We zoom in on a career as brilliant as it is nomadic—a career dedicated to the science of language.

For James Henderson it all began almost thirty years ago at MIT (the Massachusetts Institute of Technology). And he has never lost his taste for international research, working at universities including Exeter, Geneva, and Edinburgh, and also in the very heart of Xerox's research laboratories in Grenoble.

Dr. Henderson is a pioneer in his field. At the turn of the new millennium he successfully developed innovative techniques that combine neural network analysis and the understanding of natural language—one of the greatest challenges in the artificial intelligence domain.

James takes over from Professor Andrei Popescu-Belis, who has joined HEIG-VD in Yverdon. The new research group Natural Language Understanding is a continuation of Idiap's work in natural language processing, and will focus on translation and indexing and on the classification of texts and their "entailments".

Revealing opinions by using data from Facebook

A lexical or textual entailment establishes a relation between two words or two texts. For example, "this is a cat" implies that the cat is an animal, which is clearly true. But to make such links, a computer must first go through a learning process. Passionate about such questions, James wants to use the study of these linguistic entailments to "summarize" opinions. Using a model that employs entailments to extract meaning from phrases, it will be possible to build consensus from a wide variety of differing opinions. The American researcher sees a direct application in the area of opinion polls: instead of deploying questionnaires or carrying out telephone surveys, it will be possible to determine a group of people's collective opinion by extracting and analyzing their individual social media posts.

Beyond the labs—a creative state of mind

For James, Idiap offers both an intellectually stimulating environment and great freedom to researchers: "The weekly seminars, in which we present our own research, are very interesting, as are the exchanges between the doctoral and postdoctoral students of the various groups." A state of mind built on the sharing of ideas, and one that generates—according to James—healthy competition, building a community that is creative, dynamic, and encourages discovery.

Mélanie Huck and Guillaume Clivaz celebrate the first anniversary of their joining Idiap

More than a year has passed since Mélanie and Guillaume began a new adventure, joining Idiap's research and development engineering group— two young people from very different backgrounds and a group that stands out for its significant diversity.

From architecture to virtual structures

Having studied Architecture at EPFL and worked for some time in the field, Mélanie decides to embark on a new challenge: developing software in her preferred domain. She joins the School of Engineering Vaud at Yverdon to study for a further diploma. Quickly realizing that the range of opportunities open to her is much broader than she has previously imagined, she decides to change to different horizons. And that change takes her to Idiap, which she joins in February 2017. A native of Vevey who practices Tae Kwon Do at international level, Mélanie appreciates innovation and a challenge. Since her arrival, she has worked on four different and exciting projects. She loves image processing and 3-D and sees large software solutions as structures. Architecture, it seems, is never far away.

A dynamic environment in Valais

After studying Microtechnology and obtaining a Master's degree in Robotics at EPFL, Guillaume takes on two successive temporary roles, in the chemical processes industry and in carbon materials for the aluminum industry. His search for further employment leads him to Idiap and its research and development group. From day one at the Institute he works on a vast project, which leads to numerous exchanges and collaborations. But one of the young developer's goals is to immerse himself in machine learning, and in particular to tackle deep learning techniques. When not working, Guillaume also sets time aside for the region's brass band, in which he plays the saxophone, and is also a member of the local tennis club.
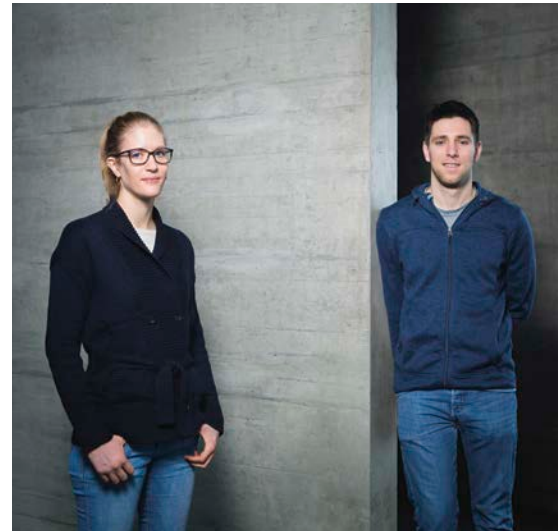
Idiap's research and development group draws its strength from the diverse interests and eclectic professional backgrounds of its engineers. These enable them to respond with originality and effectiveness to the wide variety of requests that come their way from industry and from Idiap's own researchers. These engineers also participate fully in efforts to transfer technology from research to industry—one of Idiap's three missions. And finally, it is worth noting, with pride, that many engineers who have taken on roles in Swiss or international industry come from this very team—a fact that consolidates the links that unite Idiap with the broader economic fabric. And that consolidation is, itself, a highly significant achievement.

We wish Mélanie and Guillaume a rewarding and fulfilling journey.

A jurist at the crossroads of artificial intelligence

Arriving at Idiap in March 2017, Marie-Constance Kaiflin is a veritable fountain of energy. The first jurist engaged at Idiap, she is in her element among the researchers and developers. A brilliant, modern, and necessary link between science and the law.

B efore arriving in Martigny, this native of Geneva, holder of a Master's degree in Law, had already worked in the canton. Happy with her initial experiences of the town of Sion, first at a small law firm and then in various departments of the State of Valais, she accepts an original proposal from Idiap. "Creating your own post is a great opportunity," says Marie-Constance. Propelled into the world of artificial intelligence, she is delighted to be working in an environment that is "dynamic, entrepreneurial, and cosmopolitan."

"The Swiss army knife of contract negotiation"

Her duties are as varied as they are numerous: she reviews work contracts and manages internal regulations, she analyzes and authors service contracts with industrial partners, and research conventions for projects funded by Innosuisse (the Swiss Innovation Promotion Agency). For the major European project TAPAS, coordinated by Idiap (see p. 24), she negotiates the consortium's agreements with the other eleven parties, including academic and research Institutes as well as industrial partners.

Little by little, requests for legal advice flow in, including from start-ups at the technology platform IdeArk, located in the same buildings as Idiap. She responds with the same energy to questions about contracts, about the legal constraints raised by the presence of a robot in a Finnish supermarket, and about legal research into the possible protection of a French–Swiss-German translating dictionary using copyright law.

Jurist—daily, non-stop training

How has she managed to adapt so quickly? "I have great colleagues, and I've learned a lot from them!" she exclaims. The first she mentions is Florent Monay, Head of Technology Transfer, with whom she shares an office.

She learns on-the-job, but also online. To familiarize herself with the new European legislation on data protection, she is studying the Law module of the "CAS" program in Data Protection, Privacy and Biometrics proposed by UniDistance. With the rise of digitalization and the Big Data revolution, new questions of law and ethics are emerging. And they need to be studied in a more specialized way. This is why Marie-Constance is also helping create an ethics committee. What is the secret of her success? "I'm interested in what's around me," she explains, "I like the action."

Idiap 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 2017 the Research Prize was awarded to **James Newling,** an independent and exceptionally motivated student in the field of clustering (k-means and k-medoids). The Research Publication Prize went to **Gülcan Can** for her outstanding scientific article entitled "Codical Maya Glyph Segmentation: A Crowdsourcing Approach".

External

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

8th International Joint Conference on Natural Language Processing
November 27–December 1, Taipei, Taiwan
Best Paper Award
Nikolaos Pappas and Andrei Popescu-Belis
Multilingual Hierarchical Attention Networks for Document Classification

Intelligent Service Robotics (ISR),
2017 Best Paper Award
Sylvain Calinon
A Tutorial on Task-Parameterized Movement Learning and Retrieval

20th International Conference on Artificial Intelligence and Statistics
May 2017, Fort Lauderdale, Florida, USA
Best Paper Award
James Newling and François Fleuret
A Sub-Quadratic Exact Medoid Algorithm

Gülcan Can, from Social Computing group

James Newling, from Machine Learning group

Theses completed

Five students completed their (EPFL) thesis in 2017:
Pierre-Edouard Honnet, Pedro Henrique Oliveira Pinheiro,
Olivier Canévet, Marzieh Razavi, and Gülcan Can.

Intonation Modelling for Speech Synthesis and Emphasis Preservation

Pierre-Edouard HONNET
January 2017

Thesis director: Prof. Hervé Bourlard
Members of the jury committee: Dr. Jean-Marc Vesin,
Dr. Philip N. Garner, Prof. Jean-Philippe Thiran,
Dr. Junichi Yamagishi, and Dr. Antonio Bonafonte

Large-Scale Image Segmentation with Convolutional Networks

Pedro Henrique OLIVEIRA PINHEIRO
January 2017

Thesis directors: Prof. Hervé Bourlard and
Dr. Ronan Collobert
Members of the jury committee: Prof. Jean-Philippe Thiran,
Prof. Rob Fergus, Dr. Cordelia Schmid,
and Dr. Mathieu Salzmann

Object Detection with Active Sample Harvesting

Olivier CANÉVET
February 2017

Thesis director: Dr. François Fleuret
Members of the jury commitee: Prof. Pascal Frossard,
Prof. Gilles Blanchard, Prof. Raphael Salzmann,
and Dr. Mathieu Salzmann

On Modeling the Synergy Between Acoustic and Lexical Information for Pronunciation Lexicon Development

Marzieh RAZAVI
August 2017

Thesis directors: Prof. Hervé Bourlard
and Dr. Mathew Magimai-Doss
Members of the jury committee: Prof. S. Süsstrunk,
Dr. K. Knill, Prof. M. Davel, and Prof. J.-Ph. Thiran

Visual Analysis of Maya Glyphs via Crowdsourcing and Deep Learning

Gülcan CAN
September 2017

Thesis directors: Prof. Daniel Gatica-Perez
and Dr. Jean-Marc Odobez
Members of the jury committee: Dr. Denis Gillet,
Prof. Alberto del Bimbo, Prof. Rolf Ingold,
and Prof. Jean-Philippe Thiran

Centre du Parc, Rue Marconi 19,
case postale 592, CH - 1920 Martigny

T +41 27 721 77 11
F +41 27 721 77 12
M info@idiap.ch
www.idiap.ch
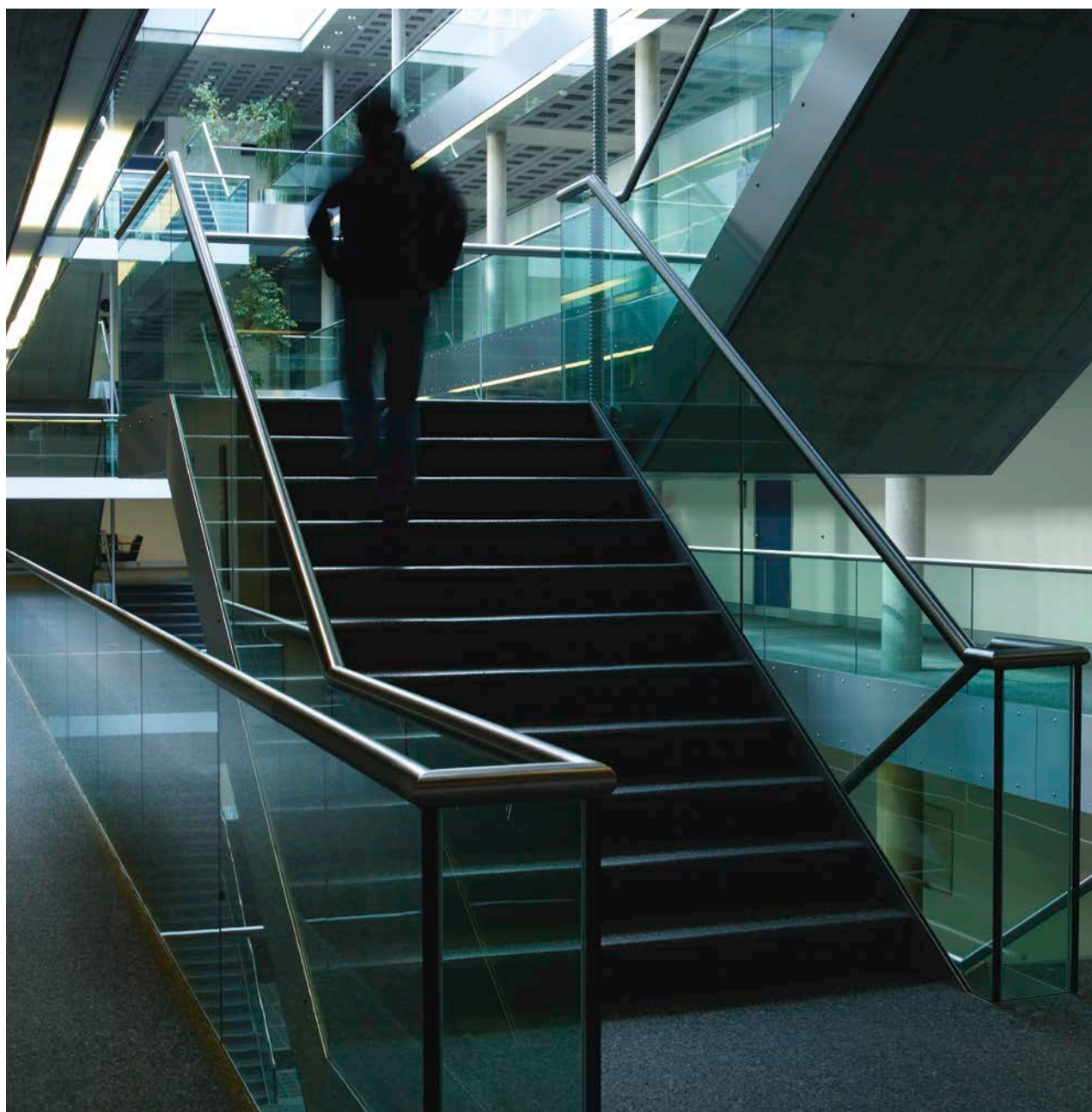
scientific inserts

I — XL

Speech and Audio Processing

Overview

Heads: Prof. Hervé Bourlard (MS and PhD, Polytechnic University, Mons, Belgium, 1982 and 1992), Dr. Philip N. Garner (MEng, University of Southampton, UK, 1991; PhD, University of East Anglia, UK, 2011), Dr. Mathew Magimai-Doss (MS by Research, Indian Institute of Technology Madras, India, 1999; PhD, Ecole Polytechnique Fédérale de Lausanne, Switzerland, 2005), Dr. Petr Motlicek (MS and PhD, Brno University of Technology, Czech Republic, 1999 and 2003).

Group overview

Speech processing has been one of the mainstays of Idiap's research portfolio for many years, covering most of the aspects of speech processing such as multilingual automatic speech recognition (ASR), speech synthesis, speech coding, speaker identification, automatic speech intelligibility evaluation, or speech processing for classification of motor speech disorders. The expertise and activities of the group encompass statistical automatic speech recognition (based on Hidden Markov Models-HMM, or hybrid systems exploiting Deep Neural Networks-DNN and new deep learning architectures), text-to-speech (TTS), speaker recognition (with extensions towards text-dependent and forensics scenarios) and generic audio processing (covering sound source localization, microphone arrays, speaker diarization, audio indexing, very low bit-rate speech coding, perceptual background noise analysis for telecommunication systems) and, more recently, Compressive Sensing (CS) and Sparse Recovering theories applied to ASR.

The Speech and Audio Processing group in 2017 was composed of 1 head of group, 3 principal investigators, 3 research associates, 6 postdocs, 9 PhD students and 4 interns.

Key scientific outputs

Idiap has always significantly contributed to both Hidden Markov Model (HMM) and Deep Neural Network (DNN) based approaches applied in acoustic modelling for various speech processing tasks. Use of techniques from HMM and HMM-DNN based speech recognition in HMM and HMM-DNN based speech synthesis resulted in a unified approach to speech recognition and synthesis. The group was well placed to take full advantage of recent advances in new architectures of deep learning, studied in particular through *PyTorch* and other open source frameworks. Advances in Automatic Speech Recognition (ASR) are usually researched through *Kaldi* toolkit, now used by most of the international speech community.

In 2017, several key contributions were achieved by the group, including: (1) multilingual speech recognition, especially in cross-lingual adaptation, and speech recognition in low-resourced language conditions, (2) speaker recognition, through both text-independent and particularly text-dependent speaker verification scenarios and information fusion for large-scale speaker identification, (3) large scale media processing, including multilingual broadcast news recognition, and spoken query for spoken term detection, and (4) development of novel deep neural network (DNN) architectures for multilingual speech recognition.

The group has also started exploiting new Compressive Sensing (CS) and Sparse Recovering theories to automatic speech recognition, developing new theoretical links between CS and statistical/HMMDNN approaches, resulting in improved ASR performance, as well as new spoken term query detection algorithms.

❉ **Additional information and a list of projects are available from** www.idiap.ch/speech.

Automatic speech recognition

In recent years, our ASR research activities have been expanded from mono-lingual to cross-/multilingual 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:
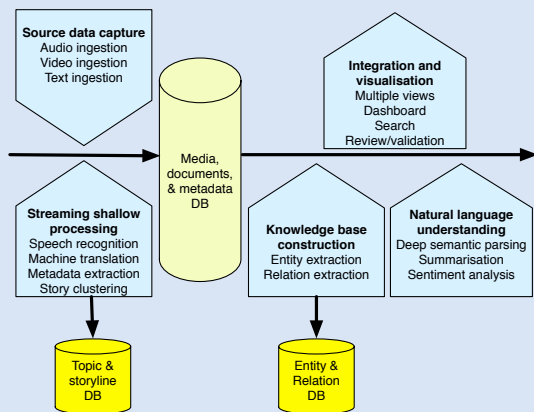
→ **Robust parametrisation and acoustic modelling**
We are still investigating new features (e.g., posterior-based features) and new acoustic models (new forms of hidden Markov models, such as KL-HMM, or artificial neural networks) that are more robust to noise and acoustic environments, as well as to speaker variability (e.g., accented speech, or dialect).

→ **Cross-lingual and multi-lingual speech recognition**
The EC SUMMA project (Scalable Understanding of Multilingual Media), as illustrated in Figure 1, aims at integrating stream-based media processing tools (including speech recognition and machine translation) with deep language understanding capabilities (including named entity relation extraction and semantic parsing), for open-source applications and implemented in use cases at the BBC and DeutscheWelle. The US IARPA SARAL

project (Summarization and domain-Adaptive Retrieval of Information Across Languages) aims at developing cross-lingual retrieval and summarization techniques that will work for any language in the world, given minimal resources to work with. In those contexts, 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 boot-strapped using well-resourced data and adapted to the target language.

German (dialect) speech recognition. A collaboration in 2017 with Swisscom enabled us to extend our Swiss German work to serverbased large vocabulary situations, and in tandem with translation to English and other Swiss languages.

Figure 1

An overview of the EC H2020 SUMMA project, a current focus for multilingual speech recognition at Idiap.
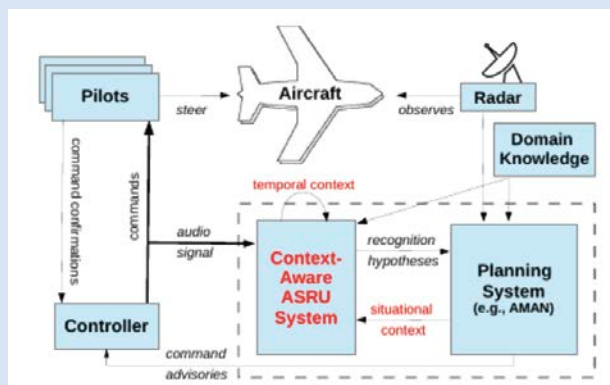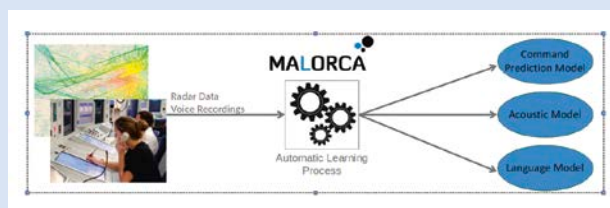
Figure 2

Up: Overview of the EC MALORCA project; semi-supervised adaptation of three models. **Down:** graphical representation of building blocks.

→ 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 also used in commercial products of the Idiap spinoff *recapp IT AG*. Work on speech recognition for Walliserdeutsch one of the most difficult to understand of the Swiss dialects, was first published in 2014; the associated database is also available for download. Since 2015, we collaborate with *recapp IT AG* on a wider range of Swiss dialects towards the first commercial product that performs Swiss

→ 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 projects A-MUSE, PHASER, and PHASER-QUAD, and CTI project MultiVEO[1]. In this context, there are currently two on-going efforts. The first focuses on development of "sparse" and "low-rank" HMM frameworks by establishing a link between statistical speech recognition formalism and the recent theories of sparse modelling and (probabilistic) principle component analysis (PCA). 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.

→ **Lexicon development**
Over the past six years, Idiap has been conducting research on the use of alternative subword units, specifically graphemes, and development of an ASR approach which is capable of handling both acoustic and lexical resource constraints by learning graphemeto-phoneme relationship through the acoustic speech signal (funded through the SNSF-project FlexASR and HASLER Foundation project AddG2SU).

→ **Semi-supervised learning of acoustic and language model parameters**
This problem has been conducted at Idiap over several past years, as it aims to develop a technology to (1) automatically select speech data which is acoustically highly informative, or (2) to automatically assess recognition output with high confidence. This type of technology is essential to allow semisupervised learning of speech recognition models using unlabelled data. One of the specific architecture has been developed within the EC H2020 MALORCA project [2], combining both acoustic (i.e. speech) and situational (i.e. radar) information to build automatic speech recognition system for air-traffic controllers (i.e. controller-pilot communication). The MALORCA project, graphically represented in Figure 2 further focused on technologies to automatically adapt the developed system for new airports, while minimising human effort to manually transcribe speech data.

→ **Punctuation prediction**
As our ASR output becomes an input for processing at higher semantic levels, for instance in the SUMMA project, other meta-data such as punctuation becomes necessary. ASR does not normally produce punctuation, but it is possible using similar techniques, notably language modelling and pause detection. More recently, this is influenced by the encoder-decoder approaches used in machine translation.

→ **Failure analysis of ASR and HSR decoding channels**
This is a novel strategy further developed in 2017 to identify the key sources of information loss in ASR and human speech recognition (HSR). This approach builds on the power of DNN in probabilistic characterization of the subword classes constituting a language. We cast ASR and HSR as separate channels decoding the sequence of subword components from their probabilistic representation. Information theoretic measures are developed to assess and quantify the information loss in acoustic modeling for ASR decoding using hidden Markov models.

Speech synthesis and coding

→ **Text-to-speech synthesis (TTS):**
Although newer than ASR, TTS is now an established venture for the speech group at Idiap. TTS has been central to several projects, including: SIWIS [3] (Spoken Interaction With Interpretation in Switzerland), D-BOX [4] and SP2 [5] (SCOPES project on speech prosody). The TTS work at Idiap was largely aligned with the statistical synthesis trend, which uses the same technology as ASR. However, the group has tracked the recent developments in deep learning which will dominate future research. SIWIS aimed to do speech to speech translation in Swiss languages, including the concept of accents. Current work under MASS (Multilingual Affective Speech Synthesis) brings the concept of emotion into the speech synthesis, particularly via modelling of prosody.

→ **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. Previously, under the RECOD project funded by Armasuisse, the technology lends itself to operate at 200–300 bits per second. The solution relies on deep learning characterization of the phone attribute features dubbed as phonological posteriors. Recent advancement building on the findings of the SNSF project PHASER, led to increased efficiency of the architecture. Unique structures of the phonological posteriors are identified as the sparse pronunciation codes composing natural speech communication; a small size codebook is thus constructed and used for ultra low-bit-rate speech coding. Moreover, work on speech coding continued under the SP2 project, mainly on aspects concerning prosody. This led to the release of the "PhonVoc" toolkit, an end-to-end neural network based phonetic and phonological vocoder.

IV

2 www.malorca-project.de

3 www.idiap.ch/project/siwis/
4 www.idiap.ch/scientific-research/projects/dbox
5 www.idiap.ch/scientific-research/projects/sp2

Speaker recognition and speech data analytics

In the context of the European SIIP project[6], illustrated by Figure 3, the Speech and Audio Processing group has significantly improved their capabilities in suspect identification applicable to very large scale data. The SIIP technology has successfully passed three proof-of-concept and field-test events in 2016 and 2017, demonstrating the performance of the developed technology among the key stakeholders. The developed suspect identification solution can analyse not only lawfully intercepted calls, but also multiple types of social-media channels. SIIP has also developed a framework allowing to combine evidences extracted by different types of engines (i.e. inter-task engines such as language/accent/gender/age identification, or keyword-spotting) to eventually improve speaker identification.

As discussed in Section 8, and as part of the SNSF Project UniTS[7], the group is also contributing to the Biometric Person Recognition group, including the development of countermeasures to detect attacks on speaker verification systems through forged speech samples.
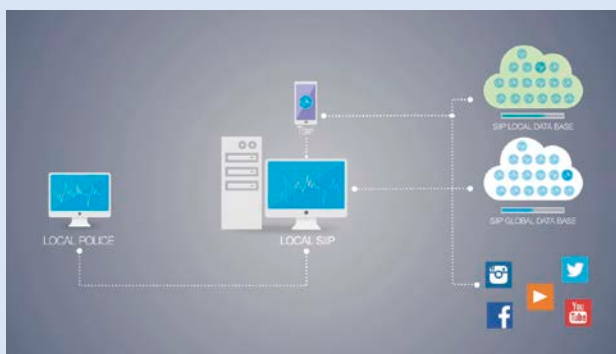
Figure 3

Illustration of SIIP speaker identification system exploiting lawfully intercepted calls, including multiple types of social-media information streams.

Forensic speaker verification and audio analysis

In 2017, Idiap continued its work in audio forensic sciences, aiming to extract relevant evidences from speech and audio signals that may be ultimately presented as admissible fact in a court of law. Besides a task of enhancement of speech recordings to improve the intelligibility, we have mostly focused on forensic voice comparison to determine the identity of the speaker associated with the strength of evidence. Idiap collaborates with various law enforcement agencies in Switzerland (e.g. Federal and cantonal Police) and abroad to provide direct support in analysing strength of evidences.

Large scale spoken query retrieval

→ **Query-by-example spoken term detection (QbE-STD) based on subspace modelling and detection**
Mainly exploiting its knowhow in Deep Neural Networks (DNN) and sparse recovery modeling, Idiap continues its research efforts towards keyword spotting and spoken term detection with a focus on searching large audio archives using spoken queries. Over the past 12 months, Idiap has developed pioneering retrieval solutions relying on the characterization and detection of the low-dimensional subspace of the DNN phonetic posteriors.

→ **Large scale spoken query indexing**
In 2017, Idiap further developed a powerful methodology for large-scale spoken query retrieval relying on hashing. Contribution of hashing is two-fold: (1) Compressing the information bearing data characteristic and (2) Identifying a data-driven symbolic space for effective search. Idiap hashing solution relies on DNN representation learning at sub-phonetic attribute level to facilitate cross-lingual applications.

→ **Objective speech intelligibility assessment**
Over the past two years, Idiap has been actively developing novel methods to objectively assess intelligibility of speech signal. This work is now exploited in the context of HASLER Foundation project FLOSS[2] and the pathological speech processing activities discussed below.

Pathological speech processing

In 2017, a new research track, focusing on the processing of pathological speech, was initiated at Idiap, mainly through two new projects:

→ MoSpeeDi (Motor Speech Disorder) SNSF Sinergia project: aiming at characterizing phonetic speech planning and motor speech programming/execution and their impairments, in collaboration with Marina Laganaro, Faculty of Psychology and Educational Science, University of Geneva, Cécile Fougeron, Laboratoire de Phonétique et Phonologie, Paris 3-Sorbonne Nouvelle, and Frédéric Assal, Neurology, Geneva University Hospitals and Faculty of Medicine, University of Geneva.

→ EC TAPAS (Training Network on Automatic Processing of PAthological Speech) coordinated by Idiap: targeting three key research problems, (1) detection (develop speech processing techniques for early detection of conditions that impact on speech production), (2) therapy (using newly-emerging speech processing techniques to produce automated speech therapy tools), and (3) assisted Living (redesigning current speech technology so that it works well for people with speech impairments and also helps in making informed clinical choices).

Sign language recognition and assessment

In the context of SNSF Sinergia project SMILE,[8] Idiap has been conducting research on sign language recognition and assessment. The consortium project coordinated by Idiap involves partners from HfH, Zurich and University of Surrey, UK. The end goal of the project is to develop a sign language assessment system that can assist Swiss German sign language learners as well as aid in standardizing a vocabulary production test that can be aligned with levels A1 and A2 of the Common European Framework of Reference for Languages (CEFR).

Sound localization and microphone array

In 2017, our research activities in the area of microphone array based speech processing and speaker diarization were mainly addressing the problem pf source localization and speech detection and also reconstruction through binary sparse coding framework. In 2017, the group has also contributed to the Perception and Activity Understanding group through EC H2020 MuMMER project, focusing on audio source localization applied in robotics.

Key publications

[1]   P.-E. Honnet, B. Gerazov, A. Gjoreski, and P. N. Garner, "Intonation modelling using a muscle model and-perceptually weighted matching pursuit", Speech Communication, Vol. 97, 2018.

[2]   D. Ram, A. Asaei, and H. Bourlard, "Sparse Subspace Modeling for Query by Example Spoken Term Detection," to be published in IEEE/ACM Transactions on Audio, Speech, and Language Processing, March 2018.

[3]   A. Asaei, C. Milos, and H. Bourlard, "Perceptual Information Loss due to Impaired Speech Production", IEEE/ACM Transactions on Audio, Speech, and Language Processing, Issue: 99, August 2017.

[4]   M. Razavi, R. Rasipuram and M. Magimai-Doss, "Towards Weakly Supervised Acoustic Subword Unit Discovery and Lexicon Development Using Hidden Markov Models", Speech Communication, Volume 96, February 2018, Pages 168–183

[5]   S. Tong, P. Garner, and Bourlard, "An Investigation of Deep Neural Networks for Multilingual Speech Recognition Training and Adaptation", Proc. of Interspeech", Proceedings of Interspeech, Stockholm, Sweden, 2017.

[6]   S. Dey, P. Motlicek, S. Madikeri, and M. Ferras, "Exploiting Sequence Information for Text-Dependent Speaker Vefirication", Proc. of 2017 IEEE Intl.Conference on Acoustics, Speech, and Signal Processing (ICASSP), New Orleans, pages 5370–5374, 2017.

Machine Learning

Overview

Head: Dr. François Fleuret (MS École Normale Supérieure and University of Paris VI, 1995; PhD, University of Paris VI, 2000; Habilitation, University of Paris XIII, 2006; EPFL MER)

Group overview

Machine learning encompasses computer techniques that aim at learning automatically representations and decision rules from data. It lies at the intersection between statistics, algorithmics, and signal processing. The main objective of the Machine Learning group is the development of novel machine-learning and statistical methods, with a strong emphasis on their algorithmic efficiency.

The application domain for the methods we develop includes image and video processing, but also industrial sensor analysis and sequence modeling.

The group was composed in 2017 of one head of group, two post-doctoral researchers, five PhD students, one intern, and it had strong interactions with the software development team. François Fleuret is also the supervisor of one phd student from the EPFL Space Engineering Center, and the co-supervisor of two phd students from the EPFL Computer Vision lab.

Key scientific outputs

In 2017, our work has resulted in contributions that improved the state-of-theart for large-scale clustering, planetary surface reconstruction from stereo images, knowledge transfer between large neural networks, training of generative adversarial models, and modeling of sequences with recurrent neural networks.

✴ **Additional information and a list of projects are available from** www.idiap.ch/ml.

Efficient Machine learning

→ **Large-scale clustering**
The k-means algorithm is one of the cornerstones of machine learning for real-world applications, in particular for large-scale training. Over the recent years, several approaches have been developed based on the triangular inequality to avoid computing distances when simple bounds insure that it is not necessary.

In the context of the MASH2 project funded by the Hasler foundation, we have first improved these exact bounds, and reduced the computational complexity of virtually all state-of-the-art methods, gaining up to a factor of two. We also have developed a novel exact algorithm which is the fastest in low dimension. In parallel, we have investigated the "batch" approaches, and proposed a new algorithm combining the triangular inequality and batches of increasing size. It resulted in a very powerful scheme that reuses computation already done over samples until statistical accuracy requires the use of additional data points.

→ **Sub-linear hard sample extraction**
Virtually every state-of-the-art detection algorithm is trained in a sequential manner, improving it iteratively with samples miss-classified by its previous version. In the SNSF project DASH, we were interested in speeding-up this "hard sample" extraction. Our approach consisted of using the prior structure provided for any large collection of images as a recursive decomposition into homogeneous sub-families to concentrate computation in a principled manner. We adapted the Monte-Carlo Tree Search to recursively apply a bandit approach to the exploration-exploitation dilemma.

In the ongoing ISUL SNSF project, we are investigating a novel approach that bypasses the need for prior knowledge, and instead relies on an efficient approximation of a sample's importance, combined with a sound estimate of the actual speed gain to expect. This allows us to activate our procedure only when the trade-off between speed and accuracy is beneficial.

→ **High-dimension similarity measures**
The SNF WILDTRACK project was a collaboration between Idiap, EPFL and ETHZ around the design of a robust multi-camera pedestrian tracking system.

One core component of such a system is a re-identification algorithm to automatically associate a person leaving and later re-entering the camera field of view. This boils down to learning a metric, so that a small distance between two images is a reliable indicator that the person is the same on both, and large values is a reliable indicator these are different persons.

The technique we have developed casts this learning as a rank optimization problem with a regularizer that enforces the linear projection to remain well-conditioned. Our method allows very large scale learning and beats all existing state-of-the-art methods on standard benchmark data-sets.
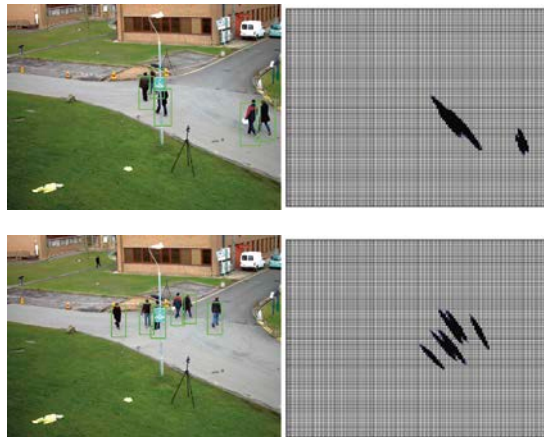
Figure 4

Person detection obtained with a deep neural network, and the corresponding occupancy map estimates.
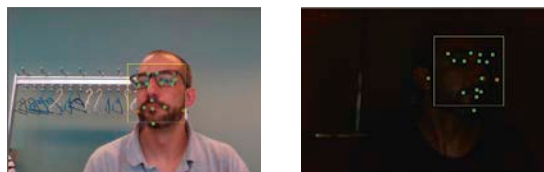
Figure 5

Face alignment examples with qualitative visibility estimation. The color from green to red indicates the visible confidence from high to low.
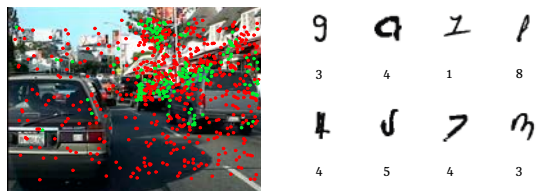
Figure 6

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 to train an object detector (left), and on difficult and unusual examples to train an image classifier (right).

Deep learning

→ **Depth estimation for planetary surface reconstruction**
In the framework of the NCCR PlanetS, and in collaboration with the EPFL eSpace center, we are developing a new algorithm to reconstruct the surface of Mars, given images provided by the ExoMars Trace Gas Orbiter. State-of-the-art methods use machine-learning to automatically match corresponding pairs of images, to estimate the depth of the surface that has been imaged. We improve upon existing technique by relaxing the need for a supervised training set, and using instead crude prior knowledge about the general topography of the planet. This will allow in the future to leverage very large unlabeled training sets.

→ **Deep learning for multi-camera detection**
We have developed in the MEMUDE project funded by the Hasler Foundation a new approach to adapt a monocular deep-learning detector to a multi-camera context. We first fine-tune it to the problem of person detection, and then fine-tune a Siamese network with one such monocular structure per view on a small multi-view data-set. This staging from a very large generic data-base to a small specific multi-view person data-set allows to beat existing state-of-the art multi-view methods.

→ **Generative Adversarial Networks**
A very promising family of methods are the so called "Generative Adversarial Networks", which rely on training jointly two models, one synthesizing realistic signals (images, sound, text) and another trying to discriminate the synthetic signal from the genuine one. Such techniques have demonstrated striking performance in many application domains, but involve a complex and unstable optimization problem.

We have developed a new method that consists of training several such pairs in parallel, and maintaining carefully their statistical independence. This insures that their joint behavior has a good "covering" property, and we show experimentally that the resulting synthesis is less likely to miss sub-families of samples.

→ **Manufacturing quality control**
High-end manufacturing requires accurate quality control, which results in a difficult computer vision problem: while "good" examples are plentiful and quite predictable, the failures are rare and diverse.

We have developed a model able to learn from the correct examples a efficient representation, specific to their appearance. The poor performance of this representation when it encodes a defective example provides an efficient detection criterion.

→ **Multi-modal accurate face localization for identification**
A standard pre-processing used to improve face recognition consists of localizing accurately face landmarks in order to register feature extraction with respect to them. In recent works, we have extended a state-of-art procedure to a multi-modal context where both a RGB image, near-infrared image, and a depth-map are available. Doing so improves performance drastically, in particular when dealing with poor lighting conditions.

VIII

→ **Low-computation Drone Detection**
In a collaboration with ArmaSuisse, we have investigated the development of efficient neural networks for drone detection in video feeds captured by in-flight drones. The main objective of this project was to adapt object detectors to the limited computation capabilities of embedded hardware, usually one order of magnitude less powerful than standard hardware for machine learning. Our approach consists in penalizing during training the use of large number of parameters, pushing the process toward the removal of certain sub-structures of the model identified as less useful in the inference.

Key publications

[1] J. Newling and F. Fleuret. K-Medoids For K-Means Seeding. In Proceedings of the international conference on Neural Information Processing Systems (NIPS), 2017.

[2] S. Tulyakov, A. Ivanov, and F. Fleuret. Weakly Supervised Learning of Deep Metrics for Stereo Reconstruction. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), pages 1348–1357, 2017.

[3] P. Baqué, F. Fleuret, and P. Fua. Deep Occlusion Reasoning for Multi-Camera Multi-Target Detection. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), pages 271–279, 2017.

[4] T. Chavdarova and F. Fleuret. Deep Multi-Camera People Detection. In Proceedings of the IEEE International Conference on Machine Learning and Applications (ICMLA), pages 848–853, 2017.

[5] J. Newling and F. Fleuret. A Sub-Quadratic Exact Medoid Algorithm. In Proceedings of the international conference on Artificial Intelligence and Statistics (AISTATS), pages 185–193, 2017. (Best paper award).

Social Computing

Overview

Head: Prof. Daniel Gatica-Perez (PhD, University of Washington, USA, 2001; EPFL Adjunct Professor)

Social computing is an interdisciplinary domain that integrates theory and models from ubiquitous computing, social media, machine learning, and social sciences to analyze human and social behavior in everyday life, and to create devices and systems that support social interaction.

The Social Computing group in 2017 was composed of one group head, four postdoctoral researchers, three PhD students, one intern, and three EPFL master students. Research lines investigated in 2017 included: social media analytics and mobile crowdsourcing for cities and health, ubiquitous conversational interaction analysis, and analysis of Maya hieroglyphic media.

→ **Key scientific outputs**
Publications on (1) social media analytics and mobile and online crowdsourcing to understand youth nightlife patterns, urban perception in cities, and eating patterns in everyday life; (2) multimodal analysis of soft skills in face-to-face employment interviews and hospitality service encounters; and (3) Maya hieroglyphic visual analysis. 13 EPFL PhD students have graduated from the group since 2002.

✳ **Additional information and a list of projects are available from** www.idiap.ch/socialcomputing.

Social media analytics and mobile crowdsourcing for cities and health

Our work in this domain spans several research lines. First, as part of the SNSF Youth@Night and Dusk2Dawn projects [9] (Multidisciplinary Study of Young People's Going Out and Drinking Behaviors, in collaboration with Addiction Switzerland and the University of Zurich), we investigated the use of mobile crowdsourcing and social media analytics to characterize urban phenomena, both in Switzerland and other world regions.[10] This included the automatic recognition of place ambiance from social media images using deep learning [1] (Figure 7), and automatic recognition of drinking activity from smartphone sensor data, for a population of 200 young volunteers in Switzerland.

Regarding mobile crowdsourcing and health, in the context of the Bites-n-Bits project [11] (Understanding Eating Routines in Context, in collaboration with Nestle Research Center), we developed a smartphone sensing app to collect everyday life data about eating and drinking for a population of 120 college students, and used it to build an approach for recognition of eating occasions [2]. Furthermore, using visual social media (Instagram) as data source, we analyzed eating and drinking patterns in the context of the Social Food and People project [12]. (supported by EPFL Integrative Food Science and Nutrition Center.)

Regarding mobile crowdsourcing for social innovation, we continued our work using the Civique platform [13] (Mobile Data for Local Causes), which included its use for teaching students about humanitarian technologies [14], and to collect multimedia data from Swiss heritage speakers living in the USA [15] (in a collaboration with the University of Lausanne, supported by the UNIL-EPFL Program on Collaborative Research on Science and Society). Finally, our work in Latin America on mobile crowdsourcing for youth engagement on local urban issues was expanded through a collaboration with institutions in Mexico and Colombia [3]. Other collaborations using the Civique platform are under development.
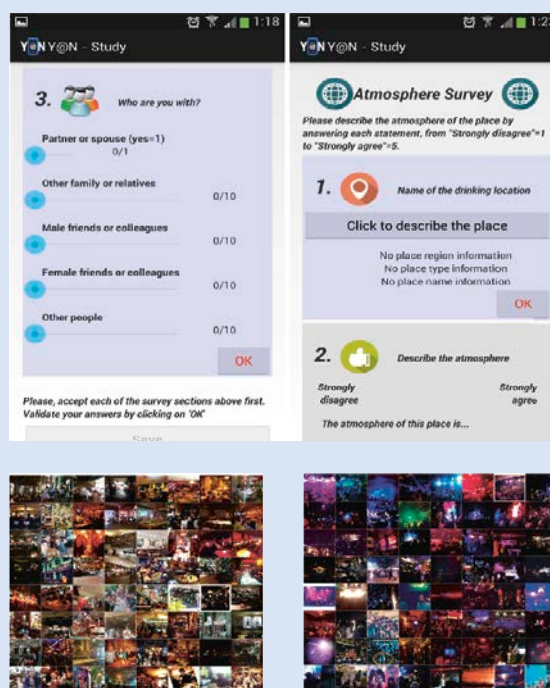
Figure 7

Up: Youth@Night survey logger app. **Down:** Social media images classified as Restaurant and Stage by convolutional neural network.

IX

9 www.youth-night.ch
10 www.youtube.com/watch?v=71ht15VAoLw
11 www.bitesnbits.org
12 www.idiap.ch/en/allnews/what-does-instagram-reveal-about-food-consumption-in-switzerland

13 www.civique.org
12 www.idiap.ch/en/allnews/
what-does-instagram-reveal-about-food-consumption-in-switzerland
14 www.idiap.ch/en/allnews/understanding-how-technology-can-revolutionize-humanitarian-work
15 wp.unil.ch/sina/

Ubiquitous interaction analytics

In the context of the SNSF UBImpressed project [16] (Ubiquitous Frist Impressions and Ubiquitous Awareness), we have developed methods to analyze dyadic interactions in the workplace using multiple sensors (cameras, Kinect devices, microphone arrays, smartphones, and wearables), and to infer variables like perceived hirability and performance (Figure 8). This is joint work with the University of Lausanne and Cornell University, in partnership with Vatel Switzerland Hotel Management School. Specifically, we found connections between automatically measured nonverbal behavior (including speaking activity, prosody, head activity, and body motion) and impressions in job interviews and hospitality service encounters [4]. Some of this work will be transfered to industry in the context of a new CTI project.

Figure 8

X   Sensors available in the social sensing lab. Kinect V2, Microcone, Q-Sensor, Android smartphone, Google Glass.

Visual analysis of Maya hieroglyphs

In the context of the SNSF MAAYA project [17] (Multimedia Analysis and Access for Documentation and Decipherment of Maya Epigraphy), we developed methods for visual analysis of hieroglyphs in ancient codices [18]. Specifically, we designed a methodology that included crowdsourced glyph localization and segmentation by non-experts for data labeling, and deep learning for visual recognition [5] (Figure 9). This included visualization of the output of deep networks for interpretation of the learned visual representations. Our work was covered by Horizons Magazine, the scientific magazine of the SNSF and the Swiss Academies of Arts and Sciences [19].

Figure 9

Left: Convex hull of the segmentation ground truth for the target glyph (red line, blue filling). Center: Gray-scale image of aggregated segmentations generated by online crowdworkers. Right: Final aggregated segmentation.

Key publications

[1]   Y. Benkhedda, D. Santani, and D. Gatica-Perez, "Venues in Social Media: Examining Ambiance Perception Through Scene Semantics," in Proc. ACM Int. Conf. on Multimedia (MM), Mountain View, Oct. 2017.

[2]   J.-I. Biel, N. Martin, D. Labbe, and D. Gatica-Perez, "Bites'n'Bits: Eating Behavior in Context from Mobile Data," PACM on Interactive, Mobile, Wearable, and Ubiquitous Technologies (IMWUT), Vol. 1, No. 4, Dec. 2017.

[3]   S. Ruiz-Correa, D. Santani, B. Ramirez Salazar, I. Ruiz Correa, F. Alba Rendon-Huerta, C. Olmos Carrillo, B. C. Sandoval Mexicano, A. H. Arcos Garcia, R. Hasimoto Beltran, and D. Gatica-Perez, "SenseCityVity: Mobile Sensing, Urban Awareness, and Collective Action in Mexico," IEEE Pervasive Computing, Special Issue on Smart Cities, pp. 44–53, Apr.-Jun. 2017.

[4]   S. Muralidhar, M. Schmid Mast, and D. Gatica-Perez, "How May I Help You? Behavior and Impressions in Hospitality Service Encounters," in Proc. ACM Int. Conf. on Multimodal Interaction (ICMI), Glasgow, Nov. 2017.

[5]   G. Can, J.-M. Odobez, and D. Gatica-Perez, "Codical Maya Glyph Segmentation: A Crowdsourcing Approach," IEEE Trans. on Multimedia, published online Sep. 2017, Vol. 20, No. 3, pp. pp. 711–725, Mar. 2018.

16 www.idiap.ch/project/ubimpressed
17 www.idiap.ch/project/maaya
18 lab.idiap.ch/maaya/

19 www.idiap.ch/en/allnews/maya-glyph-analysis-on-horizons-magazine

Perception and Activity Understanding

Overview

Head: Dr. Jean-Marc Odobez (Engineer degree, ENST Bretagne, 1990; Ms in Signal Processing, Rennes University, 1990; PhD, University of Rennes, France, 1994; EPFL MER)

Group overview

The Perception and Activity Understanding group conducts research in human activities analysis from multi-modal data. This entails the investigation of fundamental tasks like the detection and tracking of people, the estimation of their pose or the detection of non-verbal behaviors, and the temporal interpretation of this information in forms of gestures, activities, behavior or social relationships. These tasks are addressed through the design of principled algorithms extending models from computer vision, multimodal signal processing, and machine learning, in particular probabilistic graphical models and deep learning techniques. Surveillance, traffic and human behavior analysis, human-robot interactions, and multimedia content analysis are the main application domains.

In 2017, the group was composed of one group head, four post-doctoral researchers, six PhD students.

Key scientific outputs

The group is well known for its work on video sequence analysis, probabilistic tracking, non-verbal behavior extraction (in particular attention modeling), and temporal motif discovery. In 2007, the Klewel company was created using its OCR technology. In 2015 and 2016, the PAU team ranked first at the MediaEval Person discovery challenge. In 2016, its work on 3D face and gaze tracking from cheap RGB-Depth sensors has been patented and has led to the creation of the Eyeware start-up company. Besides the consolidation of his ground work, the group has investigated deep learning methods for several tasks like gesture recognition, audio-visual speaking activity modeling, gaze, joint audio localization and speech/non-speech detection, body landmark detection using depth data, and multimedia processing (cross-modal transfer learning, shape recognition, text localization). It is also working on the integration of its sensing technology and algorithms into real-time perceptual systems for human-robot interaction (Pepper platform, EU MuMMER project), or to companies (e.g. anti-tailgating detection system). During the period 2013–2017, the group published 17 journal papers and more than 40 conference papers.

✳ **Additional information and a list of projects are available from** www.idiap.ch/perception.

Deep learning

→ *Gesture Recognition.* In [5] we studied multimodal Deep Dynamic Neural Networks (DDNN) for the segmentation and recognition of short spontaneous communicative gestures, as illustrated in Fig. 10a. We adopted a semi-supervised hierarchical approach based on statistical temporal models relying on high-level spatio-temporal gesture representations trained using deep neural networks suited to the input modalities (dynamics of skeletal joint information (Fig. 10b), fusion of batches of depth and RGB images (Fig. 10c).
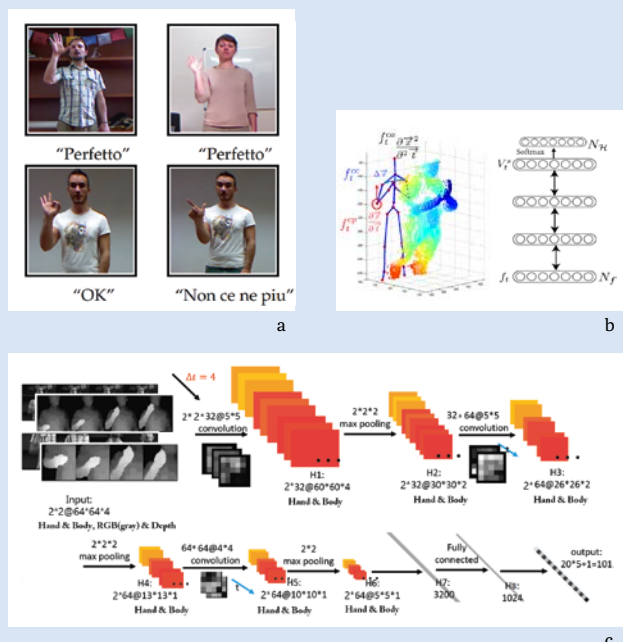
Figure 10

Multimodal gesture recognition from RGB-D data, using deep-learning methods. a) Spontaneous communicative gestures to be recognized. b) Encoding of the skeleton information with Restricted Bolzman Machines. c) CNN models of visual and depth hand data.
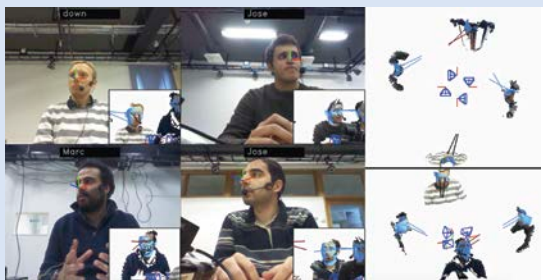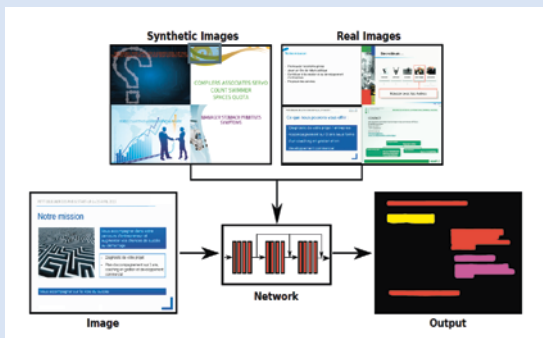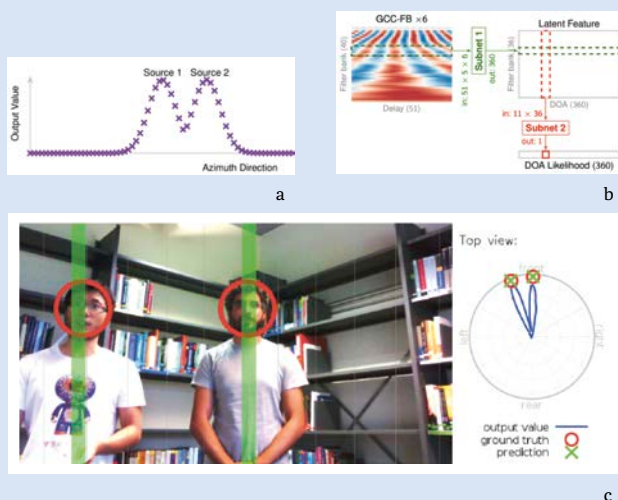
Figure 11

Up: Semantic Text Segmentation Network (STSN) trained with synthetic and real data, that apart from localizing text in images with high accuracy also classifies text regions in different semantic categories (title, bullets, standard text), and then further recognizes text (OCR). **Down:** Heads are tracked, eye localized and aligned before inferring gaze and attention towards people or other semantic labels (all information analysed in 3D).

XII

Figure 12

Multiple sound source localization for HRI. **a)** Encoding the network outpout as likelihood. **b)** Neural network architecture, mapping pairwise generalized-cross-correlation features to the direction of arrival. **c)** Sample result on real HRI data from the pepper robot.

→ *Semantic text recognition (OCR).* In the context of the CTI funded VIEW project, we investigated deep learning methods for the detection, segmentation, categorization and recognition of text content in presentations, allowing further semantic tagging (see Fig. 11). Models and systems are being transfered to the Klewel company, improving the internet referencing and visibiliy of the recorded and broadcast presentations of its client.

→ *Audio analysis.* In the context of the EU MuMMER project on social robotics, we have investigated the use of different DNN architectures for the localization of multiple speech sources, a situation that can often arise when the robot has to deal with multiple people. By encoding the potential sound source locations using a likelihood approach (Fig. 12a), exploiting a hierarchical approach (per band prediction, then fusion, Fig. 12b), the network can learn to simultaneously localize multiple sound sources (Fig. 12c). Methods were demonstrated to perform much better than state-of-the-art signal-processing based approaches (GCC or SRP-PHAT and other models or variants).

We also investigated deep-temporal representations for the modeling of audio-visual speaking activities, e.g. for differentiate dubbing from genuine talking situations for instance.

→ *Shape representation of Maya glyphs.* In the context of the SNSF MAYA project and in collaboration with the Social Computing group we investigated crowdsourcing, vizualisation and deep learning methods for the data labeling, shape representation, and glyph analysis in ancient Codices (see Fig. 9).

Multiple object tracking

Our previous works resulted in an enhanced Conditional Random Field tracking-by-detection framework with important characteristics making it successful for person or face tracking in challenging conditions: long-term pairwise links, pairwise cue measurement costs sensitive to the time interval between two detections, integration of multi-cue association reliability factors, and unsupervised learning.
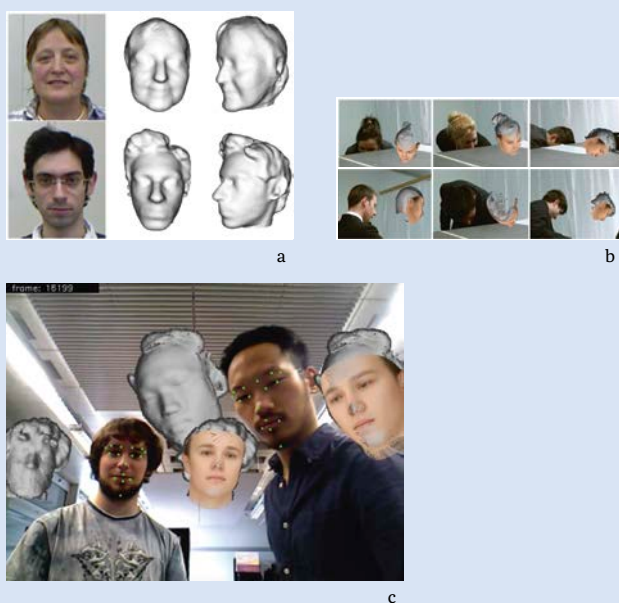
Figure 13

Robust head pose tracking. **a)** Example of automatically re-constructed heads. **b)** Tracking results in adverse conditions (UBImpressed registration desk set-up). **c)** Tracking results in an HRI situation (image taken from the Pepper robot).

Gaze, non-nerbal behavior extraction, human-robot interactions

Our long-standing work on NVB perception has been extended in several ways. In the context of the Swiss UBImpressed project, In [1] we have worked on robust and accurate head pose tracking from RGB-D data. Combining the benefits of the online fitting of a 3D face morphable model with the online 3D reconstruction of the full head (Fig. 13a), the approach provides more support when handling extreme head poses, like for the analysis of difficult registration desk sequences (SNSF Sinergia UBImpressed project, Fig. 13b), or multiple interacting people in HRI (MuMMER project, Fig. 13c).

Our long-term research on gaze estimation from cheap RGB-D (depth) cameras like Kinect [4] has been pursued. We investigated online techniques leveraging social situations to compute parameters of a gaze bias correction model. This improved gaze coding from 56% accuracy to 85% on several dozen hours of data. Figure 11 illustrates the result of applying our tools for gaze coding in multiparty situations.

We continue improving our real-time multi-person perception system (Pepper robot in the EU MuMMER project), with features including tracking, pose estimation, re-identification, extraction of non-verbal cues (attention, head gestures), as well as speech sound localization.

Multimodal face and person diarization and naming

Identifying people appearing and speaking in multimedia data as we have done in the EUMSSI EU project allows to monitor who said what and when and is thus important for the design of search and fast browsing tools of or broadcast programs. On addition to face tracking, clustering, audio-visual association, and dubbing detection, we investigated this year domain cross-modal domain adaptation and transfer learning methods for improving the estimation of short utterance speaker embedding leveraging face embeddings [3].

Key publications

[1] Y. Yu, K. Funes and J.-M Odobez. HeadFusion: 360oHead Pose tracking combining 3D Morphable Mode and 3D Reconstruction IEEE Trans. Pattern Anal. Mach. Intell., accepted for publication, 2017.

[2] G. Can, J.-M. Odobez, and D. Gatica-Perez. Codical Maya Glyph Segmentation: A Crowdsourcing Approach, IEEE Trans. on Multimedia, published online Sep. 2017, Vol. 20, No. 3, pp. pp. 711–725, Mars 2018.

[3] N. Le and J.-M. Odobez. A Domain Adaptation Approach to Improve Speaker Turn Embedding Using Face Representation, In Int. Conf. on Multimodal Interactions (ICMI), 2017.

[4] K. Funes and J.-M Odobez. Gaze estimation in the 3d space using rgb-d sensors: Towards head-pose and user invariance. Int. Journal of Computer Vision, 118(2):194–216, june 2016.

[5] D. Wu, L. Pigou, P.-J. Kindermans, N. Le, L. Shao, J. Dambre, and J.-M. Odobez. Deep dynamic neural networks for multimodal gesture segmentation and recognition, IEEE Trans. Pattern Anal. Mach. Intell., 38(8):1583–1597, 2016.

XIII

Robot Learning and Interaction

Overview

Head: Dr. Sylvain Calinon (MS and PhD, EPFL, 2003 and 2007)

Group overview

The Robot Learning and Interaction group, created in 2014, focuses on humancentric robotic applications in which the robots can learn new skills by interacting with the end-users. From a machine learning perspective, the challenge is to acquire skills from only few demonstrations and interactions, with strong generalization demands. It requires the development of intuitive active learning interfaces to acquire meaningful demonstrations, the development of models that can exploit the structure and geometry of the acquired data in an efficient way, and the development of adaptive control techniques that can exploit the learned task variations and coordination patterns.

The Robot Learning and Interaction group in 2017 was composed of one postdoctoral fellow, four PhD students, three MSc students, five BSc students and one trainee.

Key scientific outputs

Development of robot learning and adaptive control algorithms that can be applied to a wide range of applications, for robots that are either close to us (assistive robots in IDRESS), parts of us (prosthetic hands in TACT-HAND), or far away from us (manipulation skills in deep water in DexROV). Attentive to reproducible research, the group regularly releases open source codes accompanying its publications at www.idiap.ch/software/pbdlib/.

☀ **Additional information and a list of projects are available from** www.idiap.ch/rli.

Probabilistic models of movements and skills

In many robotics applications, demonstrations or experiences are sparse. In such situation, it is important to get as much information as possible from each demonstration. We explore approaches encoding demonstrations from the perspective of multiple coordinate systems. This is achieved by providing a list of observers that could potentially be relevant for the movement or the task to transfer. A statistical learning approach is then used to determine the variability and coordination patterns in the movement by considering different coordinate systems simultaneously. This approach allows the orchestration of multiple coordinate systems to reproduce movements in new situations (typically, to adapt a movement to new positions of objects). The proposed *task-parameterized* model exploits the structure of the

task, which can in many robotics problems be expressed in the form of coordinate systems or local projections. It was shown that such approach provides better generalization capability than conventional regression.
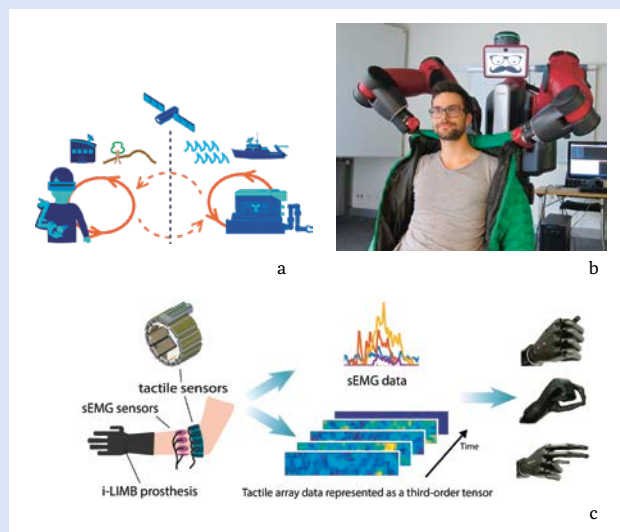
Figure 14

Human-centric robotic applications. **a)** Teleoperation assistance for the control of a bimanual robot (DexROV project), where the robot assists the teleoperator when repetitive or structured tasks are detected, as a form of human-robot collaboration. For example, if the robot observes that the task of drilling requires the drill to be perpendicular when it approaches a surface, the robot will then automatically orient the drill when it approaches another surface, letting the teleoperator concentrate on the position to drill while letting the robot maintain the correct orientation. **b)** Personalized assistance in dressing (IDRESS project), where the robot learns by demonstration the preferences and type of assistance required by each user. This is achieved by extending the movement primitives frequently employed in robotics to a wider repertoire of skills composed of reactive behaviors based on force, impedance, position and orientation. **c)** Online learning and adaptive control of a prosthetic hand (TACT-HAND project), where electromyography and tactile sensing data are combined to control a prosthetic hand.

A distinctive view of optimal control

Model predictive control (MPC) is ubiquitous in robot control, but the core formulation of this control problem and its associated algorithms can be extended to a wider range of problems, which has often been overlooked in robot learning. In particular, the most simple form of MPC (unconstrained and linear, with a homogeneous double integrator system) already has advantage for motion synthesis and planning problems, where it can be combined elegantly

with probabilistic representations of movements. This method allows the retrieval of smooth and natural trajectories analytically, by taking into account variation and coordination constraints. Instead of learning trajectories directly, the approach allows the learning of the underlying controllers to drive the robot. Namely, it learns to reject perturbations only in the directions that would affect task performance (minimal intervention control). This can typically be exploited with torque-controlled robots to regulate the tracking gain and compliance required to reproduce a task in an adaptive manner.
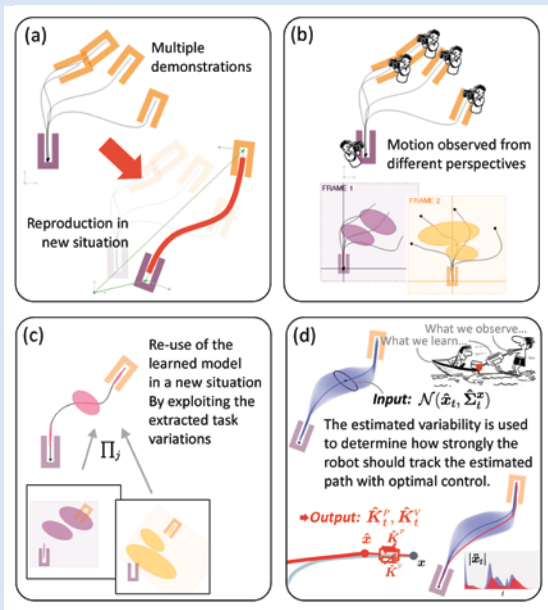

Figure 15

Task-parameterized model to synthesize movements in new situations. **a)** Observation in different situations to extract the underlying structure of the task. **b)** Probabilistic encoding of the task in multiple coordinate systems. **c)** The cross-situational observations are used to adapt the motion to new situations. **d)** Model predictive control strategy to reproduce the movement by exploiting the retrieved variability and correlation information.

Geometry-aware statistical learning and control

The data encountered in robotics are characterized by simple but varied geometries, which are often underexploited when developing learning and control algorithms. Such data range from joint angles in revolving articulations, rigid body motions, orientations represented as unit quaternions, sensory data processed as spatial covariance features, or other forms of symmetric positive definite matrices such as inertia or manipulability ellipsoids. Moreover, many applications require these data to be handled altogether.

We exploit Riemannian manifold techniques to extend algorithms initially developed for Euclidean data, by efficiently taking into account prior knowledge about these manifolds and by modeling joint distributions among these heterogeneous data. The use of these differential geometry

techniques allow us to treat data of various forms in a unified manner (including data in standard Euclidean spaces). It can typically be used to revisit common optimization problems in robotics formulated in standard Euclidean spaces, by treating them as unconstrained problems inherently taking into account the geometry of the data.
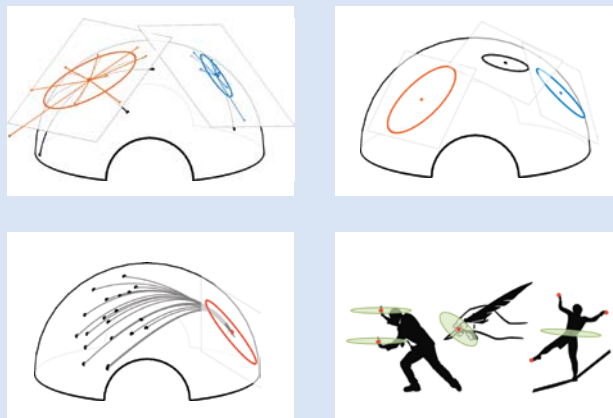

Figure 16

Statistics and control on Riemannian manifold. **From left to right and top to bottom:** Encoding as Gaussian mixture model, fusion of information with product of Gaussians, linear quadratic tracking, and example of application consisting of learning and tracking manipulability ellipsoids (symmetric positive definite matrices), which are used to find poses that a robot should adopt to face external perturbations.

Key publications

[1]   Pignat, E. and Calinon, S. (2017). Learning adaptive dressing assistance from human demonstration. Robotics and Autonomous Systems, 93, 61–75.

[2]   Jaquier, N., Connan, M., Castellini, C. and Calinon, S. (2017). Combining Electromyography and Tactile Myography to Improve Hand and Wrist Activity Detection in Prostheses. Technologies, 5:4, Special Issue on Assistive Robotics.

[3]   Zeestraten, M.J.A., Havoutis, I., Silvério, J., Calinon, S. and Caldwell, D.G. (2017). An Approach for Imitation Learning on Riemannian Manifolds. IEEE Robotics and Automation Letters (RA-L), 2:3, 1240–1247.

[4]   Calinon, S. (2016). A Tutorial on Task-Parameterized Movement Learning and Retrieval. Intelligent Service Robotics, 9:1, 1–29.

[5]   Berio, D., Calinon, S. and Leymarie, F.F. (2017). Generating Calligraphic Trajectories with Model Predictive Control. In Proc. of the 43rd Conf. on Graphics Interface, pp. 132–139.

Uncertainty Quantification and Optimal Design

Overview

Head: Prof. David Ginsbourger (Ph.D. Mines Saint-Etienne 2009, Habilitation Universität Bern 2014, Titularprofessor Universität Bern 2018)

Group overview

The Uncertainty Quantification and Optimal Design group focuses on quantifying and reducing uncertainties in the context of natural and artificial complex systems. Application domains notably include energy and geosciences, with a number of collaborations ranging from safety engineering to hydrology and climate sciences. In all these fields the study of complex systems often relies on expensive data acquisition and model runs, calling for adapted experimental design strategies.

UQOD started at Idiap in September 2015, with members coming from and keeping strong academic ties to the Institute of Mathematical Statistics and Actuarial Science (IMSV) of the University of Bern (UniBE). During the year 2017, the UQOD group has been composed of a permanent senior researcher, an Idiap postdoctoral researcher (now at IDSIA), an intern, and two UniBE students occasionally visiting the group (one master student and a PhD student who was in "co-tutelle" with Marseille).

Key scientific outputs

Current contributions include efficient algorithms for Bayesian set estimation, notably for estimating and quantifying uncertainties on overcritical parameter regions with Gaussian Process (GP) models. In 2017, the project "Learning and visualizing dangerous regions in multivariate parameter spaces" funded by the Hasler foundation has been conducted within the UQOD group. Other recent results deal with the interplay between covariance kernels and properties of associated GPs, with implications in function prediction under structural constraints and beyond. Ongoing work also encompasses novel algorithms for non-stationary modelling of extremes with application in climate sciences, contributions to statistical modelling of hail, as well as collaborations with geoscientists in uncertainty quantification and simulation-based optimization.

✳ **Additional information and a list of projects are available from** www.idiap.ch/uqod.

Bayesian optimization and emulation with Gaussian Process models

Bayesian global optimization relying on GPs has become a standard for optimizing prohibitively expensive to evaluate systems, e.g. with response(s) of interest stemming from heavy numerical simulations. It constitutes one of the core domains of expertise of the UQOD group, with recent contributions ranging from theoretical to methodological questions (such as parallelization, handling large data sets, coping with high-dimensional inputs, etc.) and applications. Notably, the group has been involved in a collaboration with researchers in hydrogeology from the universities of Lausanne and Neuchâtel, with the aim to investigate Bayesian optimization for contaminant source localization relying on flow simulations. The preprint "Contaminant source localization via Bayesian global optimization" (currently available online as a Hydrology and Earth System Sciences discussion paper) illustrates some first optimization results based on a fixed geology: Bayesian optimization automatically selects the ground truth contamination source much more efficiently than brute force grid search. This ongoing work has also lead to generate objective functions that are publicly available for algorithm benchmarking.

The UQOD group also investigates GP emulation per se, with a particular focus on the incorporation of expert knowledge and the identification of structural properties of objective functions through the specification of covariance kernels and the estimation of their parameters.

Recent work includes non-stationary GP modelling using warping approaches and also connections to multiresolution analysis and wavelet modelling, see for instance [3]. Ongoing work (presented at the ISI World Statistic Congress 2017) also concerns the estimation of weights by maximum likelihood in multiple kernel Gaussian Process modelling.

Computer experiments for the quantification and the reduction of uncertainties

Besides global optimization, UQOD has also been focusing on strategies dedicated to other goals such as locating parameter regions leading to a response exceeding a given threshold, corresponding e.g. to an abnormal behaviour of the considered system.

In the project "Learning and visualizing dangerous regions in multivariate parameter spaces", novel approaches based on profile extrema have been developed for visualizing and quantifying uncertainties on excursion regions with application to a coastal flooding test case provided by colleagues from BRGM (Bureau de Recherches Géologiques et Minières, Orléans, France).

Figure 17 represents the kind of output delivered by the developed approach: for each individual parameter of interest, maximum/minimum responses with respect to all remaining parameters are estimated (with confidence statements) in function the individual parameter in question.
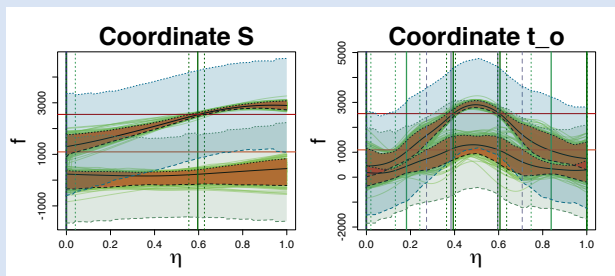
Figure 17

Estimating profile optima on a coastal flooding BRGM application

More generally, the combination of GP modelling and stochastic simulation techniques have been investigated within UQOD for quantifying and reducing uncertainties on sets, with example applications in safety engineering, where the typical goal is to identify the set of dangerous and/or safe configurations of a complex system and if possible to provide some measures of confidence along with the estimate(s).

Motivated by the estimation of such sets, a research work has been conducted leading to a class of algorithms called "asymmetric nested Monte Carlo", that prove useful for computing conservative estimates of excursion sets and more generally for efficiently estimating orthant probabilities of high-dimensional Gaussian vectors [2].

These results have been used to derive conservative set estimates on a neutronic criticality safety test case (from IRSN, the French Institut de Radioprotection et de Sûreté Nucléaire) starting from a initial experiments, and also to derive sequential design strategies dedicated to this class of conservative set estimation problem, that were applied to this IRSN test case and also to another BRGM test case. This ongoing work, in collaboration with colleagues from Neuchâtel, CentraleSupélec and IRSN, is summarized in the preprint "Adaptive Design of Experiments for Conservative Estimation of Excursion Sets" now available online (arXiv:1611.07256).

UQOD has also been active in a collaboration with colleagues from CentraleSupélec and Toulouse University towards establishing convergence results (See arXiv:1608.01118) for some important classes of Sequential Uncertainty Reduction strategies, with links to Bayesian optimization.

Statistical analysis of non-stationary time series arising in climate sciences

Extreme value analysis is a key approach to a number of phenomena from nature, economic and industry activities. The theory of extreme value distributions has been studied for decades, and adequate statistical distributions are known for dealing with maxima and quantify how severe and with what probability extreme events may occur. However, many methods in use require a sample of independent identically distributed maxima in order for the underlying mathematical theory to be valid. This poses serious challenges to climate scientists and beyond, as phenomena exhibit dependencies and also change over time. In approaches that have been recently developed within UQOD and in collaboration with the Oeschger Center for Climate Change Research, e.g. within the Internship of Alan Maître, the main focus has been on extreme value analysis of non-stationary climatological times series.
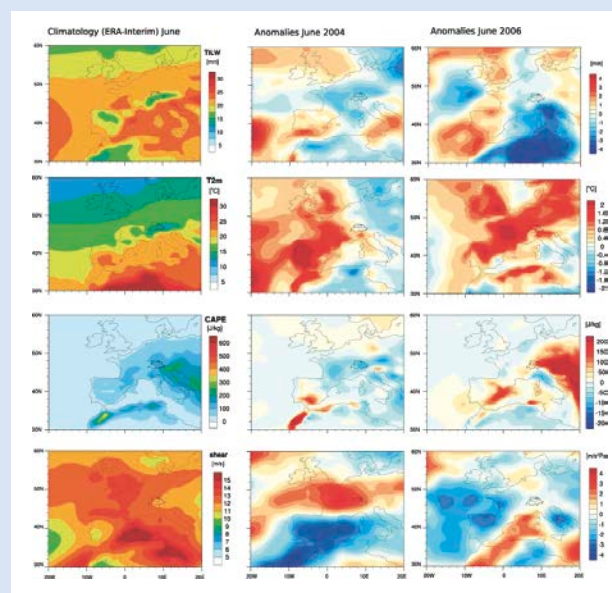
Figure 18

Environmental variables and their anomalies, used for hail prediction

Also, UQOD has been involved in statistical modelling of monthly numbers of hail days based on monthly averaged, large-scale environmental variables. A Poisson regression model coping with non-stationarity via a month categorical variable and continuous anomalies of environmental variables (See Figure 18) enabled producing useful predictions [1] despite the loss of information occasioned by focusing on monthly quantities and neglecting higher-frequency phenomena.

Computational Bioimaging

Overview

Head: Dr. Michael Liebling (MS, EPFL, 2000; PhD, EPFL 2004; postdoc, Caltech, 2004–2007; Assistant Prof (2007–2013), Associate Prof (2013–2017), Adjunct Prof (2017–) University of California Santa Barbara (UCSB)

Group overview

Research in the Computational Bioimaging Group focuses on developing image acquisition, reconstruction and analysis algorithms to study live biological systems. Practical tools aim at (i) extending the physical limits of imaging hardware via techniques including super-resolution and multi-view, space variant deconvolution, and (ii) quantitative analysis of complex biological systems: motion-based image analysis, cell tracking, microscopic fluid flow estimation, and integration of multimodality images.

In 2017, the Computational Bioimaging Group was composed of the head of group and three PhD students.

Key scientific outputs

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

❈ **Additional information and a list of projects are available from** www.idiap.ch/cbi.

Direct inversion algorithm for focal plane scanning optical projection tomography

To achieve approximately parallel projection geometry, traditional optical projection tomography (OPT) requires the use of low numerical aperture (NA) objectives, which have a long depth-of-field at the expense of poor lateral resolution. Particularly promising methods to improve spatial resolution include ad-hoc post-processing filters that limit the effect of the systems' modulation transfer function (MTF) and focal-plane-scanning OPT (FPS-OPT), an alternative acquisition procedure that allows the use of higher NA objectives by limiting the effect of their shallow depth of field yet still assumes parallel projection rays during reconstruction. We have established the existence of a direct inversion formula for FPS-OPT and, based on this formula, we have proposed a

Key publications

[1] A Poisson regression approach to model monthly hail occurrence in Northern Switzerland using large-scale environmental variables, Erica Madonna, David Ginsbourger and Olivia Martius, in: Atmospheric Research, 203:261–274, 2018.

[2] Estimating orthant probabilities of high dimensional Gaussian vectors with an application to set estimation, Dario Azzimonti and David Ginsbourger, in: Journal of Computational and Graphical Statistics, Accepted author version posted online: 03 Aug 2017.

[3] Non-parametric warping via local scale estimation for non-stationary Gaussian process modelling, Sébastien Marmin, Jean Baccou, Jacques Liandrat and David Ginsbourger, in: Wavelets and Sparsity XVII, International Society for Optics and Photonics, 2017.

[4] On uncertainty quantification in hydrogeology and hydrogeophysics, Niklas Linde, David Ginsbourger, James Irving, Fabio Nobile and Arnaud Doucet, in: Advances in Water Resources, 110:166–181, 2017.

point spread function-aware algorithm that is similar in form and complexity to traditional filtered backprojection (FBP). With simulations, we have demonstrated that the point spread-function aware FBP for FPS-OPT leads to more accurate images than both traditional OPT with deconvolution and FPS-OPT with naive FBP reconstruction. We further illustrated the technique on experimental zebrafish data, which showed that our approach reduces out-of-focus blur compared to a direct FBP reconstruction with FPS-OPT.

Figure 19

We used focal plane scanning OPT (FPS-OPT) to image the head of a Tg(*fli1a*:eGFP) zebrafish in 3D fluorescence with a 10×/0.3NA air objective. Under such conditions, single-plane OPT would be unable to produce an acceptable reconstruction due to the large sample thickness and shallow depth-of-field. With FPS-OPT, we compare 3D reconstructions from **a)** standard filtered backprojection (FBP) and **b)** our proposed PSF-aware FBP. Our proposed PSF-aware FBP algorithm reconstructs an image with less out-of-focus blur. Scale bar is 100 μm. Adapted from [1].

Platform for Reproducible Acquisition, Processing, and Sharing of Dynamic, Multi-Modal Data

In this project, a joint effort between the Computational Bioimaging group, the Robotics and Interaction group and the Biometrics Security and Privacy group, we assembled a platform for reproducible acquisition, processing, and sharing of dynamic, multi-modal data. This modular platform consists of four layers: sensing, positioning, computing, and storage. It accepts a wide range of sensing devices and complements Idiap's already existing infrastructure for collecting data in the areas of computer vision, biometry, and speech processing. Concretely the platform consists of the following components (Figure 20):

→ Two 7-axis backdrivable robotic arms (Panda, Franka Emika), computer-driven

→ Pneumatic gripper system

→ Two-color Open-SPIM light-sheet microscope (20-water immersion microscope object, 488 et 561 nm laser excitation for green and red fluorescence imaging), computer-driven

→ Dedicated data processing, storage and serving backend (4GPU NVidia K80 (19968 GPU cores), 8CPU Intel (112 CPU cores), 40TB storage)

The platform is connected to Idiap's computing and storage infrastructure, with platform-dedicated computing and storage nodes that were integrated within Idiap's computing facility. During setup and calibration, we have used the platform to acquire preliminary evaluation data. In particular, we adjusted the microscope's 3D sample positioning stage and acquired image series of fixed samples (Figs. 20). We foresee that the platform will enable new research projects both within Idiap and with outside research groups in Valais, Switzerland, and beyond.
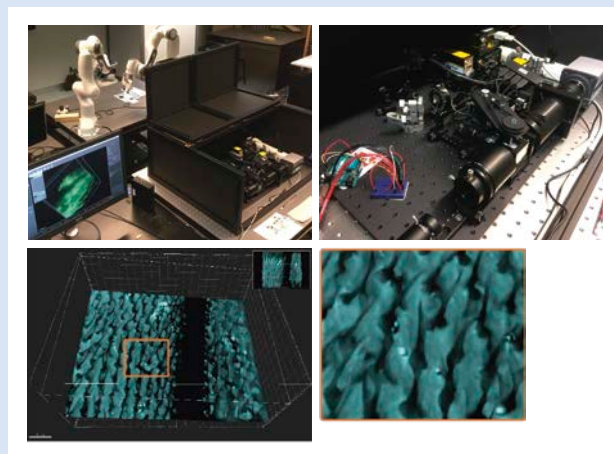
Figure 20

Photograph of the reproducible acquisition platform. Two-color Open-SPIM microscope in foreground, with robotic arms (Panda, Franka Emika, GmbH, Munich, Germany) in background; in addition to the data acquisition platform visible here, the platform relies on dedicated data processing, storage and serving backend. Right two frames, Demonstration data (Wing of a common housefly) acquired with microscopy platform. (data: Christian Jacques). Width of detail rectangle is 100 μm.

Key publications

[1]   Kevin G. Chan and Michael Liebling, "Direct inversion algorithm for focal plane scanning optical projection tomography," Biomed. Optics Express, Vol. 8, Issue 11, pp. 5349–5358 (2017).

Biometrics Security and Privacy

Overview

Head: Dr. Sébastien Marcel (PhD, University of Rennes, France, 2000; Visiting Professor, University of Cagliari, 2010; Lecturer, EPFL, 2013-)

Group overview

Biometrics refers to the automatic recognition of individuals based on their physiological and/or behavioral characteristics. The Biometrics Security and Privacy (BSP) group at Idiap focuses on three main areas of research:

→ Biometric recognition: We investigate and develop new biometrics based recognition algorithms, notably for face, voice, and vein biometric modalities.

→ Presentation attack detection: We look for new and better ways of detecting presentation attacks on face, voice, and vein biometric recognition systems.

→ Biometric template protection: We research effective methods of preserving both the security of biometric recognition systems and the privacy of their users by protecting the biometric models ("templates") that are employed by the system for recognition purposes.

The BSP group prioritizes reproducibility in research. This is important for ensuring that our work can be both verified and built upon by the wider research community. To enable reproducibility, we mainly make use of our Python-based signal-processing and machine-learning toolbox, Bob (http://www.idiap.ch/software/bob/), which we make freely available for academic purposes. The group also develops and maintains the BEAT platform (https://www.beat-eu.org/platform/) – a MLaaS platform compliant with Swiss and European data-security norms.

The group participates in several large-scale biometrics projects at Swiss (SNSF), European (H2020) or world-wide levels (eg. IARPA/DARPA) but also conducts projects directly with companies.

The BSP group provides also expertise to the Swiss Center for Biometrics Research and Testing, which, among other things, carries out tests and evaluations on commercial products related to biometrics.

The BSP group in 2017 was composed of 1 head of group, 4 research associates, 5 postdocs, 3 PhD students, 1 intern and 1 engineer.

Key scientific outputs

The BSP group has been pioneering the work on mobile biometrics (face and speaker recognition) and on PAD in face

and speaker recognition by sharing the first open databases, organising the first International competitions and producing the first reproducible research studies in the domain. Regarding face PAD in mobile scenario, the group confirmed that the current trend using discriminant classifiers is prone to over-fitting hence resulting in a lack of generalisation on unseen presentation attacks. Regarding voice PAD we demonstrated that the existing methods generalise poorly when different databases or different types of attacks are used for training and testing. These results question the efficiency and practicality of the existing PAD systems, as well as, call for creation of databases with larger variety of realistic speech presentation attacks. The BSP group also investigated approaches for heterogeneous face recognition and vein recognition.

☀ **Additional information and a list of projects are available from** www.idiap.ch/biometric.

Face and speaker recognition

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

Figure 21

Illustration of the heterogeneous face recognition problem: matching Visible spectra and Near-Infrared spectra **(left)**, matching Visible spectra and sketch **(middle)**, matching Visible spectra and Thermal spectra **(right).**

xx

Heterogeneous face recognition

The task of Heterogeneous Face Recognition (Figure 21) consists in to match face images that were sensed in different modalities, such as sketches to photographs, thermal images to photographs or near infrared to photographs. We we proposed a novel and generic approach based on Intersession Variability Modelling to handle this task.

Presentation attack detection

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 presentation attacks (aka spoofing attacks): an action of outwitting a biometric sensor by presenting a counterfeit biometric evidence of a valid user (Figure 22). 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 presentation attacks: a biometric system is usually designed to only recognise 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 presentation attack detection (PAD aka anti-spoofing) schemes for them has attracted much less attention.

Speaker recognition (SR) systems are highly vulnerable to presentation attacks (Figure 23) limiting their wide practical deployment. Therefore, to protect against such attacks, effective PAD techniques, need to be developed. We focused on the integration of PAD and SR systems, demonstrating the effect of score-based integration on recognition and attack detection performance.
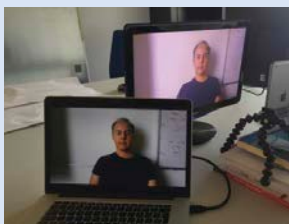
Figure 22

Illustration of video and audio presentation attacks.
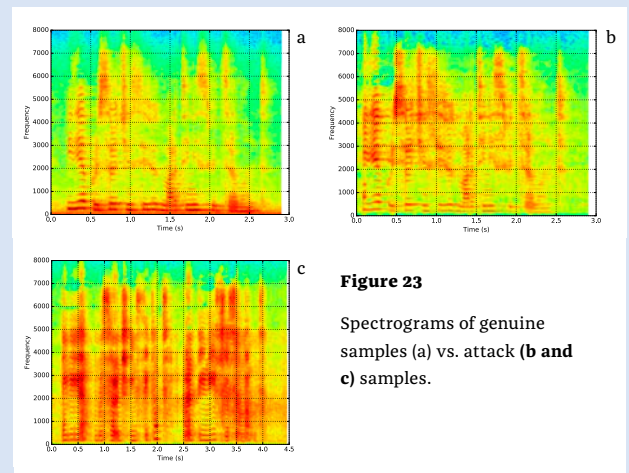
Figure 23

Spectrograms of genuine samples (a) vs. attack **(b and c)** samples.

Remote photoplethysmography

Photoplethysmography (PPG) consists in measuring the variation in volume inside a tissue, using a light source. The aim of remote photoplethysmography (rPPG) is to measure the same variations, but using ambient light instead of structured light and widely available sensors such as a simple webcam (Figure 24).

XXI
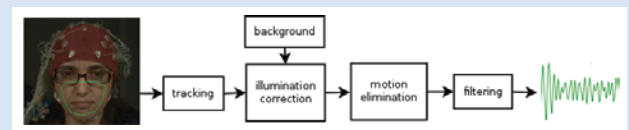
Figure 24

Illustration of remote photoplethysmography: colors from the video signal are filtered to produce an estimation of the heart beat signal.

We presented a new, publicly available database containing a relatively large number of subjects recorded under two different lighting conditions. Also, three state-of-the-art rPPG algorithms from the literature were selected, implemented and released as open source free software.

Vein biometrics

Vein recognition relies on the complex network of blood vessels located under the human skin. The vascular image of veins, that are located about 3 mm below the surface of the skin, is typically captured under near infra-red (NIR, wavelength between 700 and 1000 nm) illumination. The vein pattern can then be extracted with image pre-processing techniques (Figure 25) and used for recognition by any pattern recognition method.

Figure 25

Illustration of image processing for vein biometrics (first row for data subject 1 and second row for data subject 2): the original image (left) is geometrically normalised (middle left) and binarised with maximum curvature (middle right) or repeated line tracking (right).
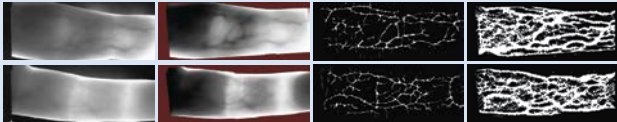
Swiss Centre for Biometrics Research and Testing

XXII In 2014, the Idiap Research Institute launched the "Swiss Centre for Biometrics Research and Testing" (www.biometrics-center.ch), a competence centre within the Institute following recent successes in coordinating International research projects in Biometrics (MOBIO, TABULA RASA and BEAT). The aim of this centre is to serve as a legacy for these projects and to push for industry-driven research and testing in biometrics.

The centre attracted the attention of large companies (license, research and testing agreements) and led to many new projects (DARPA, IARPA, CTI). In 2017, the centre has developed over 3 directions:

→ Improving the security of the BEAT platform: an external IT security company conducted an audit to identify vulnerabilities of the infrastructure, allowing our engineers to strengthen the platform.

→ Enriching the BEAT platform: our researchers and engineers developed Deep Learning and speech processing technologies to enrich the BEAT platform with new content. More particularly our engineers developed a new feature to execute GPU simulations at the heart of modern AI tools.

→ Testing: our researchers conducted the independent testing of biometrics products from SMEs (e.g., KeyLemon now acquired by AMS) and the evaluation of a prototype for a large corporation.

Key publications

[1] A. Mohammadi, S. Bhattacharjee and S. Marcel, "Deeply Vulnerable – a study of the robustness of face recognition to presentation attacks", IET Biometrics, 1–13, 2017.

[2] P. Korshunov and S. Marcel, "Impact of score fusion on voice biometrics and presentation attack detection in cross-database evaluations", IEEE Journal of Selected Topics in Signal Processing, 695–705, 2017.

[3] H. Muckenhirn, P. Korshunov, M. Magimai-Doss and S. Marcel, "Long-Term Spectral Statistics for Voice Presentation Attack Detection", IEEE/ACM Transactions on Audio, Speech and Language Processing, 2017.

[4] S. Bhattacharjee and S. Marcel, "What you can't see can help you: extended-range imaging for 3D-mask presentation attack detection", IEEE International Conference of the Biometrics Special Interest Group (BIOSIG), 2017.

[5] A. Anjos, L. El-Shafey and S. Marcel, "BEAT: An Open-Science Web Platform", International Conference on Machine Learning, 2017.

Natural Language Processing

Overview

Head: Dr. Andrei Popescu-Belis (MEng, École Polytechnique, France, 1995; MSc, Université Pierre et Marie Curie, France, 1996; PhD, Université Paris-Sud, France, 1999)

Group overview

The Natural Language Processing group studies text analysis at the semantic and pragmatic levels for two main applications: machine translation (MT) and information retrieval (IR). The group aims to improve the state of the art on core semantic and pragmatic analysis problems, such as the disambiguation of nouns and pronouns, topic labeling, keyword extraction, or sentiment analysis – so that they improve in turn MT and IR. Regarding MT, we combine text-level processing techniques with statistical and neural MT systems to improve translation quality. Regarding IR, we design models of document content and sentiment to improve multimedia classification and recommendation.

In 2017, the NLP group had the following members: the head of the group, one postdoctoral student, two PhD students, and two interns. In September 2017, the head of the group started a professorship at HEIG-VD / HES-SO, while a new senior researcher was recruited in the same field. Therefore, the NLP group became the NLU group, which is covered in a different section of this report.

Key scientific outputs

In 2017, we have demonstrated that the semantic analysis of noun phrases is also beneficial to neural MT (in additional to statistical phrase-based MT). We have shown that word sense disambiguation could be combined with statistical and neural MT, and that coreference resolution was helpful to statistical MT, using source/target coreference similarity as a component of the objective function. The hierarchical neural networks with attention proposed for cross-lingual transfer on a document classification in 8 languages have met with considerable success, earning a best paper award at one of the main NLP conferences (IJCNLP 2017, Taiwan).

Document-level machine translation

We focus on the translation of words that depend on long-range relationships between sentences. The main examples are discourse connectives (which have been our main focus

in 2010–2014),[20] verb tenses (studied in 2013–2014), and pronouns and noun phrases (our current focus).[21] The NLP group coordinates a consortium with teams from Geneva, Zürich and Utrecht, which has reached an end in 2017. Our main 2017 achievements in document-level MT were the following ones. We designed a method to enforce the consistency of noun translation, as illustrated in Fig. 26. The method is based on a classifier that decides whether two occurrences of the same noun in a source text should be translated consistently. We combined these classifiers with phrase-based SMT for Chineseto- English and German-to-English, in two ways: automatic post-editing of noun translations vs. re-ranking of MT hypotheses. Our method closes up to 50% of the gap in BLEU scores between the baseline and the oracle classifier.

In addition to noun consistency, we proposed two methods to improve the translation of pronouns based on coreference links. First, we built a coreference model, which captures the probabilistic connection between a pronoun and the features of its antecedent (gender, number and humanness) by learning from the output of an anaphora resolution system on parallel texts. When used in combination with a phrase-based MT decoder, our method improves pronoun translation for English-French and Spanish-English MT. Second, we generalized our studies and started to improve the translation of entity mentions (nouns or pronouns) with the help of coreference resolution. We implemented a coreference-aware translation system that helps to disambiguate the translation of mentions by optimizing the similarity of mention-grouping in source and target documents, either by re-ranking entire sentence, or by post-editing mentions only. Again, this improved Spanish-English translation of pronominal mentions. Finally, using results from the DiscoMT 2015 shared task, we showed that a reference-based metric for the translation of pronouns correlates well with human judgments.

Figure 26

Inconsistent translations of repeated nouns, in blue, from German **(Ex. 1)** and Chinese **(Ex. 2)** into English. Only Ex. 1 is truly mistaken, and is corrected by our MT system.

Text analysis for multilingual document representations

We develop methods for text and multimedia information retrieval, classification and recommendation, based on the analysis of content and sentiment, and on the networked structure of (social) media repositories. This work was supported by the SUMMA EU project[22].

Within a weighted multiple-instance framework, we showed how to jointly learn to focus on relevant parts of documents and to classify them into categories. The model can operate on various sentence and document representations, including intermediate hidden states of a neural network, and has state-of-the-art results for aspect rating prediction (i.e. multi-aspect sentiment analysis). Generalizing our previous work on sentiment analysis, we demonstrated the explanatory power of our model by comparing its predicted aspect saliency values with those assigned by humans, and we showed that the aspect saliencies benefit text segmentation and summarization. We also showed that this model is mathematically equivalent to attention mechanisms recently proposed for neural networks.

We developed a state-of-the-art architecture for multilingual document modeling, based on hierarchical neural networks. We explored 3 options for multilingual training over disjoint label spaces: sharing the parameters at each layer of the network, sharing the attention mechanisms at each layer, and sharing both. We demonstrated that these architectures are useful for transferring knowledge to low-resource languages, and that they also improve over monolingual models on the full-scale data. The models were applied to topic labeling in the SUMMA project, using 600k articles from Deutsche Welle, in 8 languages, with multiple topic labels. We thus delivered a deep multilingual tagging library for the SUMMA demonstrator, illustrated in Fig. 27 for one test article. This work received the Best Paper Award at the IJCNLP 2017 Conference in Taiwan.

XXIV

Key publications

[1] Pappas N. & Popescu-Belis A. (2017) – Multilingual Hierarchical Attention Networks for Document Classification. Proceedings of the 8th International Joint Conference on Natural Language Translation (WMT), Copenhagen, Denmark. Processing (IJCNLP), Taipei, Taiwan. Best paper award.

[2] Pu X., Pappas N. & Popescu-Belis A. (2017) – Sense-Aware Statistical Machine Translation using Adaptive Context-Dependent Clustering. Proceedings of the 2nd Conference on Machine Translation (WMT), Copenhagen, Denmark.

[3] Pu X., Mascarell L. & Popescu-Belis A. (2017) – Consistent Translation of Repeated Nouns using Syntactic and Semantic Cues. Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL), Valencia, Spain.

[4] MiculicichWerlen L. & and Popescu-Belis A. (2017) – Using Coreference Links to Improve Spanishto-English Machine Translation. Proceedings of the EACL Workshop on Coreference Resolution Beyond OntoNotes (CORBON), Valencia, Spain.

[5] Pappas N. & Popescu-Belis A. (2017) – Explicit Document Modeling through Weighted Multiple-Instance Learning. Journal of Artificial Intelligence Research (JAIR), vol. 58, p. 591–626.
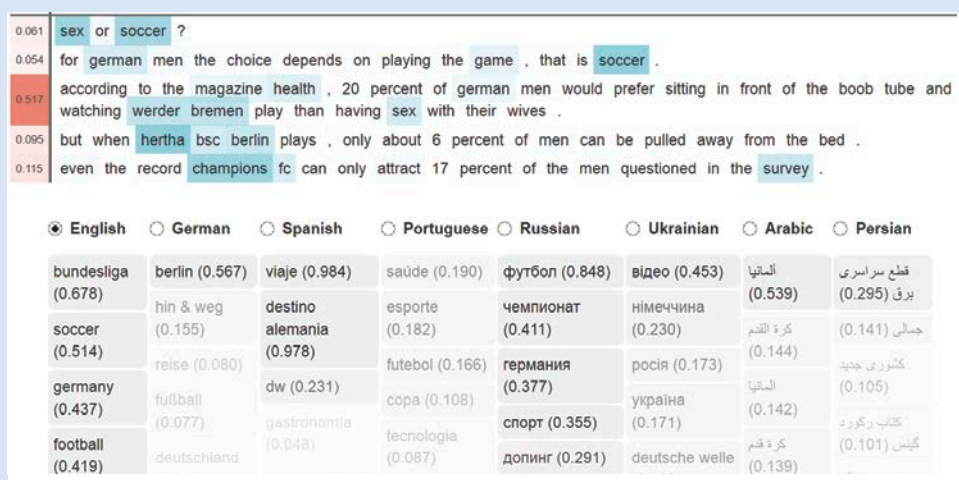
Figure 27

Multilingual topic labeling in 8 languages of a news article from Deutsche Welle.

Natural Language Understanding

Overview

Head: Dr. James Henderson (BSc, Massachusetts Inst. Technology, USA, 1987; MSE & PhD, Univ. Pennsylvania, USA, 1991,1994; MER & Chargé de Cours, Univ. Geneva, 2008–2012,2012–)

Group overview

The Natural Language Understanding group was created in September 2017, in part as a continuation of the previous Natural Language Processing group. The NLU group studies deep learning for natural language processing tasks, focusing on models with learned representations of the meaning of text. These tasks include machine translation, information retrieval, language modelling, syntactic and semantic parsing, sentiment analysis, text classification and entailment detection, applied to both text and speech. We focus on recurrent and attention-based neural network models, vector-space representations for entailment (rather than similarity), unsupervised learning of semantic representations, and multi-task learning.

From September until December 2017, the NLU group had the following members: the head of the group, one postdoctoral student, and two PhD students.

Key scientific outputs

From September to December 2017 the work of the NLU group has included both work started within the previous Natural Language Processing group and work developing new lines of research. Notable in the former is work showing improvements in neural machine translation by adding an attention mechanism over the previous target-side words, and label-aware text classification, where text descriptions of output classes are used in deep learning models to not only generalise to classes with no training data ("zero-shot" learning) but, for the first time, to improve accuracies on classes that have training data. The new lines of research build on recent results by the head of group on the unsupervised learning of word embeddings that predict entailment. Two grant proposals were submitted that develop this entailment-vectors framework for sentence-level textual entailment and its application in opinion summarisation and the analysis of the interpretation of official announcements.

Neural Network Architectures for NLP Tasks

The impressive initial results from using deep learning architectures in NLP tasks, such as for neural machine translation, are now being improved by modifying the architectures to better embody inductive biases that are important for NLP tasks. This is manifested in our work on self-attention in the decoder for neural machine translation. Instead of generating the translation with the standard LSTM decoder, we add an attention mechanism over the prefix of generated words. This better models non-local correlations in the output sentence. We find that this mechanism works best if the attention applies directly to the word embeddings, rather than the LSTM states at each word.

Another way to import inductive biases into a deep learning architecture is to explicitly model the similarity between output classes. This can be done with what we call label-aware models, where the text of the output class label is used to compute a vector for each output class, such that similarity between classes is represented as similarity between their vectors. Previously such models have generalised well in "zero-shot" learning, where the output classes at test time do not occur at all in the training data. With our proposed use of cross-entropy loss in our attention-based label-aware architecture for text classification, we showed for the first time that such models can also show improvement for the output classes which do occur in the training data, where previously non-label-aware classification has performed better. This effect is particularly marked with very large label sets, where our model also our model also trains faster, in time independent of the label set size.

Vector-Space Models of Entailment

Entailment is the relation which reflects adding information or abstracting away from information, and is fundamental to many theories of semantics. But typically vector-space models, such as word embeddings and deep learning architectures, are based on similarity between vectors, not entailment. Prior to joining Idiap, the head of group has developed a vector-space framework for modelling entailment. Unprecedented results have been achieved on the unsupervised learning of word embeddings that predict entailment between words (hyponymy), using a novel entailment-based distributional semantic model. This framework extends to entailment between larger texts (textual entailment), and has particular relevance to large-scale summarisation tasks. Textual entailment is important for summarisation because the summary must abstract away from unimportant information but still contain as much information as possible from the text.

This line of work has been the basis of two grant proposals submitted at Idiap in 2017. The first grant proposal includes exploiting the entailment-vectors framework and the word embedding results to build models of the fundamental problem of textual entailment. In parallel, the project would apply textual entailment to developing models and resources for large-scale opinion summarisation. Opinion summarisation is where a large collection of texts must be summarised, including the popularity of different opinions. The second grant proposal is for an inter-disciplinary project investigating the interpretation of official announcements. The interpretation text is a form of summary of the announcement text, so textual entailment is again relevant. But typically interpretation involves more complex reasoning and background knowledge, making it a challenging testbed.

Key publications

[1]   Lesly Miculicich Werlen, Nikolaos Pappas, Dhananjay Ram & Andrei Popescu-Belis (to appear) – Self-Attentive Residual Decoder for Neural Machine Translation. Proceedings of the 16th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL 2018), New Orleans, Louisiana, USA.

Active and Granted Projects in 2017

An overview of the projects that have been active during the year 2017 is presented below. The projects are grouped by funding agencies, starting with European and International Agencies, then Swiss Agencies to finish with the projects funded by industrial partners. Finally the list of projects accepted during 2017 but starting in 2018 is presented.

Projects in Progress during 2017

Projects Funded by European and International Agencies

Name	**MACADAMS** (Modifying Adhoc Centralised Advertisement with Digit Arena Multicast over Satellite)
Funding	Eurostars Program
Coordinator	Digit Arena
Duration	2016.09.01 – 2018.08.30
Partner(s)	Idiap Research Institute, Ecole Centrale de Lyon, Eurovision (European Broadcast Union)

Name	**SIIP** (Speaker Identification Integrated Project)
Funding	FP7
Coordinator	Verint System Ltd
Duration	2014.05.01 – 2018.04.30
Partner(s)	Idiap Research Institute, SAil Labs Technology AG, Singularlogic Anonymos etairia Pliroforiakon Systimaton & Efarmogon Pliroforikis, Green Fusion Limited (Data Fusion International), Synthema S.R.L., Ok2Go, Loquendo Spa, International Biometric Group (Uk) Limited, Cassidian SAs, Rijksuniversiteit Groningen, Inov Inesc Inovacao – Instituto de Novas Tecnologias, University of Warwick, Laboratorio Di Scienze della Cittadinanza, The International Criminal Police Organization, Police Service of Northern Ireland, Ministério Da Justiça, Lisboa

Name	**DEXROV** (Effective Dexterous ROV Operations in Presence of Communications Latencies)
Funding	H2020-BG
Coordinator	Space Applications Services
Duration	2015.03.01 – 2018.08.31
Partner(s)	Idiap Research Institute, Jacobs University Bremen, Comex SA., Interuniversity Center of Integrated Systems For The Marine Environment, Graal Tech S.R.L.

Name	**TESLA** (An Adaptive Trust-based e-assesment System for Learning)
Funding	H2020-ICT
Coordinator	Fundacio Per A La Universitat Oberta de Catalunya
Duration	2016.01.01 – 2018.12.31
Partner(s)	Idiap Research Institute, Imperial College London, The Open University, Télécom Sudparis, Open Universiteit Nederland, European Association For Quality Assurance In Higher Education Aisbl, Universite de Namur Asbl, AGència Per A La Qualitat del Sistema Universitari de Catalunya, Lplus Gmbh, Sofiiski Universitet Sveti Kliment Ohridski, Protos Sistemas de Información, S.L., Technical University of Sofia, Anadolu University, Jyvaskylan Yliopisto, European Quality Assurance Network For Informatics Education E.V., Instituto Nacional de Astrofisica, Optica Y Electronica, Wfsw, SA, Institut Mines-Telecom

Name	**MUMMER** (MultiModal Mall Entertainment Robot)
Funding	H2020-ICT
Coordinator	University of Glasgow
Duration	2016.03.01 – 2020.02.28
Partner(s)	Idiap Research Institute, Centre National de La Recherche Scientifique, Aldebaran Robotics, Teknologian Tutkimuskeskus Vtt, Kiinteistö Oy Ideapark Ab

Name	**SUMMA** (Scalable Understanding of Multilingual Media)
Funding	H2020-ICT
Coordinator	University of Edinburgh
Duration	2016.02.01 – 2019.01.31
Partner(s)	Idiap Research Institute, University College London, British Broadcasting Corporation, deutsche Welle, Priberam Informatica S.A., Leta, Qatar Computing Research Institute

Name   **4DHEART** (4D analysis of heart development and regeneration using advanced light microscopy)
Funding   H2020-MSCA
Coordinator   Fundacion Centro Nacional de Investigaciones Cardiovasculares Carlos Iii
Duration   2016.10.01 – 2020.09.30
Partner(s)   Idiap Research Institute, Centre National de La Recherche Scientifique, Universität Bern, Acquifer AG, Bitplane AG, Leica Microsystems Cms Gmbh, 4D-Nature ImAGing Consulting, S. L., Centre Europeen de Recherche En Biologie et Medecine

Name   **TAPAS** (Training Network on Automatic Processing of PAthological Speech)
Funding   H2020-MSCA
Coordinator   Idiap Research Institute
Duration   2017.11.01 – 2021.10.31
Partner(s)   University of Sheffield, Philips, Radboud University Nijmegen – Stichting Katholieke Universiteit, Ludwig-Maximilians-Universität München, Institut de Recherche En Informatique de Toulouse, Antwerpen University Hospital, Friedrich-Alexander-Universität Erlangen Nuernberg, Instituto de Engenharia de Sistemas E Computadores, Investigacao E desenvolvi-mento Em Lisboa, Interuniversitair Micro-Electronica Centrum Imec Vzw, Stichting Het Nederlands Kanker Instituut – Antoni Van Leeuwenhoek Ziekenhuis, Universitaet Augsburg

XXVIII

Name   **MALORCA** (Machine Learning of Speech Recognition Models for Controller Assistance)
Funding   H2020-SESAR
Coordinator   Deutsches Zentrum Fuer Luft – Und Raumfahrt Ev
Duration   2016.04.01 – 2018.03.31
Partner(s)   Idiap Research Institute, Universität des SAarlandes, Rízení Letového Provozu Ceské Republiky, Austro Control Osterreichische Gesellschaft Fur Zivilluftfahrt Mbh

Name   **SWAN** (Secure Access Control over Wide Area Network)
Funding   Research Council of Norway
Coordinator   Hogskolen I Gjovik
Duration   2016.01.01 – 2019.12.31
Partner(s)   Idiap Research Institute, Morpho, Bankenverband, Universitetet I Oslo, Zwipe As

Name   SAVI (Spotting Audio-Visual Inconsistencies)
Funding   USA DARPA
Coordinator   Sri International
Duration   2016.05.19 – 2020.05.18
Partner(s)   Idiap Research Institute

Name   **BATL** (Biometric Authentification with Timeless Learner)
Funding   USA IARPA
Coordinator   University of Southern California
Duration   2017.03.01 – 2021.02.28
Partner(s)   Idiap Research Institute

Name   **SARAL** (Summarization and domain-Adaptive Retrieval of Information Across Languages)
Funding   USA IARPA
Coordinator   University of Southern California
Duration   2017.10.01 – 2019.03.31
Partner(s)   Idiap Research Institute, Massachusetts Institute of Technology, Raytheon Company, Reenselaer Polytechnic Institute, University of Massachusetts Amherst, Northeastern University

Name   **SAVI** (Spotting Audio-Visual Inconsistencies)
Funding   USA DARPA
Coordinator   Sri International
Duration   2016.05.19 – 2020.05.18
Partner(s)   Idiap Research Institute

Projects Funded by Swiss Agencies

Name   **SOCIALFOOD** (Social Food and People: Characterizing Food Consumption Behaviour on Visual Social Media)
Funding   EPFL Collaboration
Coordinator   Idiap Research Institute
Duration   2017.01.01 – 2017.12.31
Partner(s)   École Polytechnique Fédérale de Lausanne

Name   **MULTI-CROWD** (Tracing Swiss Heritage Speakers' Identities in North America)
Funding   EPFL Collaboration
Coordinator   University of Lausanne
Duration   2017.02.01 – 2018.02.28
Partner(s)   Idiap Research Institute

Name   **CREM-IDIAP** (Pour une recherche fondamen-tale et appliquée au service des systèmes énergétiques territoriaux en Valais)
Funding   Etat du Valais
Coordinator   Centre de Recherches Energetiques et Municipales
Duration   2016.11.01 – 2017.03.31
Partner(s)   Idiap Research Institute

Name   **ADDG2SU-EXT.** (Flexible Acoustic data-driven Grapheme to Subword Unit Conversion)
Funding   Hasler Foundation
Coordinator   Idiap Research Institute
Duration   2016.03.01 – 2017.02.28

Name   **DCROWDLENS** (Crowdsourcing the Assessment of Deep Visual Explanations for Subjective Variables)
Funding   Hasler Foundation
Coordinator   Idiap Research Institute
Duration   2017.05.15 – 2018.01.15

Name   **FLOSS** (Flexible Linguistically-guided Objective Speech aSessment)
Funding   Hasler Foundation
Coordinator   Idiap Research Institute
Duration   2017.03.01 – 2020.02.29

Name **MASH-2** (Massive Sets of Heuristics for Machine Learning II)
Funding Hasler Foundation
Coordinator Idiap Research Institute
Duration 2013.07.01 – 2017.09.30

Name **MEMUDE** (Multi-view Detection with Metric-learning for Deep Network Fusion)
Funding Hasler Foundation
Coordinator Idiap Research Institute
Duration 2017.06.01 – 2019.05.31

Name **SETVISU** (Learning and visualizing dangerous regions in multivariate parameter spaces)
Funding Hasler Foundation
Coordinator Idiap Research Institute
Duration 2016.12.01 – 2017.08.31

Name **GPUS** (Acquisition d'un cluster de calcul dédié à l'Intelligence Artificielle)
Funding Loterie Romande
Coordinator Idiap Research Institute
Duration 2017.09.01 – 2018.08.31

Name **VALAIS+** (Une plateforme pour mieux connaître l'espace de vie du canton)
Funding Loterie Romande
Coordinator Idiap Research Institute
Duration 2015.09.01 – 2017.03.31

Name **ODESSA** (Online Diarization Enhanced by re-cent Speaker identification and Sequential learning Approaches)
Funding SNF – ANR
Coordinator Centre National de La Recherche Scientifique
Duration 2016.03.01 – 2019.02.28
Partner(s) Idiap Research Institute, Eurecom

Name **LIVEHEART** (The Cellular Basis of Cardiac Development Revealead by Live Imaging)
Funding SNF – ANR
Coordinator Institut de Génétique et de Biologie Moléculaire et Cellulaire
Duration 2016.06.01 – 2019.05.31
Partner(s) Idiap Research Institute, University of Bern, École Polytechnique Paris

Name **TACT-HAND** (Improving control of prosthetic hands using tactile sensors and realistic machine learning)
Funding SNF – DACH
Coordinator deutsches Zentrum Fuer Luft – Und Raumfahrt Ev
Duration 2016.04.01 – 2019.03.31
Partner(s) Idiap Research Institute, Universitaet Bielefeld

Name **MAAYA** (Multimedia Analysis and Access for Documentation and Decipherment of Maya Epigraphy)
Funding SNF – DACH
Coordinator Idiap Research Institute
Duration 2013.09.01 – 2017.04.30
Partner(s) University of Geneva, University of Bonn

Name **COMETS-M** (Computational Methods for Temporal Super-resolution Microscopy)
Funding SNF – Division II
Coordinator Idiap Research Institute
Duration 2016.04.01 – 2019.03.30

Name **HFACE** (HeterogeneousFace Recognition)
Funding SNF – Division II
Coordinator Idiap Research Institute
Duration 2014.07.01 – 2018.06.30

Name **ISUL** (Importance sampling for large-scale unsupervised learning)
Funding SNF – Division II
Coordinator Idiap Research Institute
Duration 2017.03.01 – 2020.02.28

Name **MASS** (Multilingual Affective Speech Synthesis)
Funding SNF – Division II
Coordinator Idiap Research Institute
Duration 2017.05.01 – 2020.04.30

Name **PHASER-QUAD** (Parsimonious Hierarchical Automatic Speech Recognition and Query Detection)
Funding SNF – Division II
Coordinator Idiap Research Institute
Duration 2016.10.01 – 2018.09.30

Name **UNITS** (Unified Speech Processing Framework for Trustworthy Speaker Recognition)
Funding SNF – Division II
Coordinator Idiap Research Institute
Duration 2015.07.01 – 2018.06.30

Name **I-DRESS** (Assistive Interactive robotic system for support in DRESSing)
Funding SNF – ERA NET
Coordinator Idiap Research Institute
Duration 2015.12.01 – 2018.11.30

Name **PLATFORM-MMD** (Platform for Reproducible Acquisition, Processing, and Sharing of Dynamic, Multi-Modal Data)
Funding SNF – R'equip
Coordinator Idiap Research Institute
Duration 2016.07.01 – 2017.12.31

Name **WILDTRACK** (Tracking in the Wild)
Funding SNF – Sinergia
Coordinator Idiap Research Institute
Duration 2014.01.01 – 2017.12.31
Partner(s) École Polytechnique Fédérale de Lausanne, Eidgenoessische Technische Hochschule Zuerich

Name **MOSPEEDI** (Motor Speech Disorders)
Funding SNF – Sinergia
Coordinator University of Geneva
Duration 2017.10.01 – 2020.09.30
Partner(s) Idiap Research Institute, University Hospitals of Geneva, Université Paris 3

Name	**YOUTH@NIGHT** (A multi-disciplinary multi-method study of young people's outgoing and drinking behaviors)
Funding	SNF – Sinergia
Coordinator	Sucht Schweiz – Research Institute
Duration	2014.01.01 – 2017.12.31
Partner(s)	Idiap Research Institute, University of Zurich

Name	**UBIMPRESSED** (Ubiquitous First Impressions and Ubiquitous Awareness)
Funding	SNF – Sinergia
Coordinator	Idiap Research Institute
Duration	2014.01.01 – 2017.12.31
Partner(s)	Université de Neuchâtel, Cornell University

Name	**MODERN** (Modeling discourse entities and relations for coherent machine translation)
Funding	SNF – Sinergia
Coordinator	Idiap Research Institute
Duration	2013.08.01 – 2017.07.31
Partner(s)	University of Geneva, Universiteit Utrecht, University of Zurich

Name	**SMILE** (Scalable Multimodal sign language Technology for sIgn language Learning and assessmEnt)
Funding	SNF – Sinergia
Coordinator	Idiap Research Institute
Duration	2016.03.01 – 2019.02.28
Partner(s)	University of Surrey, University of Applied Sciences of Special Needs Education

xxx

Name	**DUSK2DAWN** (Characterizing Youth Nightlife Spaces, Activities, and Drinks)
Funding	SNF – Sinergia
Coordinator	Idiap Research Institute
Duration	2017.07.01 – 2019.12.31
Partner(s)	University of Zurich, La Trobe University

Name	**ALLIES** (Autonomous Lifelong learnIng intelLigent Systems) Funding SNF -ERA-NET
Coordinator	Idiap Research Institute
Duration	2017.12.01 – 2020.11.30
Partner(s)	Laboratoire national de métrologie et d'essais, Université du Maine, Universitat Politecnica de Catalunya

Name	**DEEPCHARISMA** (Deep Learning Charisma) Funding UNIL Collaboration
Coordinator	Idiap Research Institute
Duration	2017.04.01 – 2017.12.31

Projects Funded by Industrial Partners

Name	**CUAV** (Feasibility Study "Countering Mini-UAVs with Mini-UAVs")
Funding	ArmaSuisse
Coordinator	Idiap Research Institute
Duration	2017.06.01 – 2017.08.31

Name	**CAMPRO** (Camera de profondeur)
Funding	Fondation The Ark
Coordinator	Idiap Research Institute
Duration	2017.08.01 – 2018.01.31
Partner(s)	Digit Arena

Name	**RECOMEDIND** (RecoMed Industrialisation)
Funding	Fondation The Ark
Coordinator	Recapp
Duration	2017.04.01 – 2017.08.31
Partner(s)	Idiap Research Institute, Clinique Romande de Réadaptation

Name	**ELEARNING-VALAIS-3.0** (eLearning-Valais 3.0)
Funding	Fondation The Ark
Coordinator	Formation Universitaire à Distance
Duration	2016.03.01 – 2018.01.31
Partner(s)	Idiap Research Institute, Klewel

Name	**LIFE** (PdG-fatigue)
Funding	Fondation The Ark
Coordinator	Clinique Romande de Réadaptation
Duration	2016.08.01 – 2017.03.31
Partner(s)	Idiap Research Institute, Life

Name	**COBALT** (Content Based Call Filtering)
Funding	Fondation The Ark
Coordinator	Idiap Research Institute
Duration	2017.07.01 – 2018.06.30
Partner(s)	Katia SA

Name	**NMTBENCHMARK** (Training and Benchmarking Neural MT and ASR Systems for Swiss Languages)
Funding	Industrial Project
Coordinator	Idiap Research Institute
Duration	2017.01.01 – 2017.12.31
Partner(s)	École Polytechnique Fédérale de LauSAnne

Name	**SNACK** (Bites'n'Bits: Understanding Eating Routines in Context)
Funding	Industrial Project
Coordinator	École Polytechnique Fédérale de Lausanne
Duration	2016.03.01 – 2018.05.31
Partner(s)	Idiap Research Institute

Name	**BBC-BOB** (Adapting bespoke speaker identification models to be used in IDIAP's 'Bob. Spear' speaker recognition software)
Funding	Industrial Project
Coordinator	Idiap Research Institute
Duration	2017.05.01 – 2017.06.30

Name	**SWISKO** (Swiss-Korean project to develop and integrate new wearable sensors into the existing DomoSafety ambient sensor system.)
Funding	CTI
Coordinator	Idiap Research Institute
Duration	2016.01.01 – 2018.02.28
Partner(s)	DomoSafety S.A., University of Bern, Hes-So Vaud

Name	**REGENN** (Robust Eye-Gaze Estimation Deep Neural Network)
Funding	CTI
Coordinator	Idiap Research Institute
Duration	2017.09.01 – 2018.12.31
Partner(s)	Eyeware

Name	**UNICITY** (3D scene understanding through machine learning to secure entrance zones)
Funding	CTI
Coordinator	Idiap Research Institute
Duration	2017.03.01 – 2018.11.30
Partner(s)	Hes-So Fribourg, Fastcom Technology SA

Name	**IMIM** (Intelligent Monitoring for In-line Manufacturing)
Funding	CTI
Coordinator	AiSA Automation Industrielle SA
Duration	2016.07.01 – 2018.05.31
Partner(s)	Idiap Research Institute

Name	**BIOWAVE** (A BIOmetric Watch Activated by VEins)
Funding	CTI
Coordinator	Centre Suisse D'Electronique et de Microtechnique
Duration	2016.02.01 – 2017.12.31
Partner(s)	Idiap Research Institute, Biowatch SA

Name	**3DFINGERVEIN** (3D FingerVein Biometrics)
Funding	CTI
Coordinator	Hes-So Valais
Duration	2016.06.01 – 2018.03.31
Partner(s)	Idiap Research Institute, Itservices Sonna Sàrl

Name	**VIEW-2** (Visibility Improvement for Events Webcasting)
Funding	CTI
Coordinator	Hes-So Valais
Duration	2016.08.01 – 2018.03.31
Partner(s)	Idiap Research Institute, Klewel

Name	**FARGO** (Convenient and Secure 3D Face Recognition based on RGB-D Cameras)
Funding	CTI
Coordinator	Idiap Research Institute
Duration	2016.05.01 – 2018.02.28
Partner(s)	Keylemon SA

Name	**ESGEM** (Enhanced Swiss German mEdia Monitoring)
Funding	CTI
Coordinator	Idiap Research Institute
Duration	2016.06.01 – 2018.01.31
Partner(s)	Recapp, Argus der Presse AG

Name	**MULTIVEO** (High Accuracy Speaker-Independent Multilingual Automatic Speech Recognition System)
Funding	CTI
Coordinator	Idiap Research Institute
Duration	2013.11.01 – 2018.03.31
Partner(s)	Veovox

Projects Awarded in 2017 and Starting in 2018

Name	**SHAPED** (Speech Hybrid Analytics Platform for consumer and Enterprise Devices)
Funding	CTI
Coordinator	Idiap Research Institute
Duration	2018.03.01 – 2020.02.29
Partner(s)	Logitech S.A.

Name	**RISE** (Rich Interpersonal Skill analytics for rEcruitment)
Funding	CTI
Coordinator	Idiap Research Institute
Duration	2018.05.01 – 2020.04.30
Partner(s)	University of Lausanne

Name	**MEMMO** (Memory of Motion)
Funding	H2020-RIA-ICT
Coordinator	Centre national de la recherche scientifique
Duration	2018.01.01 – 2021.12.31
Partner(s)	Idiap Research Institute, University of Edinburgh, Max Planck Society for the Advancement of Sciences, University of Oxford, PAL ROBOTICS SL, AIRBUS SAS, Wandercraft, Centre de médecine physique et de réadaptation, Costain Group PLC

Name	**MPM** (Multimodal People Monitoring)
Funding	Idiap-CSEM Program
Coordinator	Idiap Research Institute
Duration	2018.02.01 – 2019.01.31
Partner(s)	Centre Suisse d'Electronique et de Microtechnique

Name	**RAPPS** (Reinforced audio processing via physiological signals)
Funding	Idiap-CSEM Program
Coordinator	Idiap Research Institute
Duration	2018.03.01 – 2019.02.28
Partner(s)	Centre Suisse d'Electronique et de Microtechnique

Name	**DOMAT** (On-demand Knowledge for Document-level Machine Translation) Funding SNF – Division II
Coordinator	Idiap Research Institute
Duration	2018.01.01 – 2018.12.31
Partner(s)	HES-SO Vaud

Name	**ROSALIS** (Robot skills acquisition through active learning and social interaction strategies)
Funding	SNF – Division II
Coordinator	Idiap Research Institute
Duration	2018.04.01 – 2022.03.31

Name	**SHISSM** (Sparse and hierarchical Structures for Speech Modeling)
Funding	SNF – Division II
Coordinator	Idiap Research Institute
Duration	2018.01.01 – 2021.12.31

Articles in scientific journal 2017

[1] Erica Madonna, David Ginsbourger and Olivia Martius, A Poisson regression approach to model monthly hail occurrence in Northern Switzerland using large-scale environmental variables, in: Atmospheric Research, 2017

[2] Marzieh Razavi and Mathew Magimai.-Doss, A Posterior-Based Multi-Stream Formulation for G2P Conversion, in: IEEE Signal Processing Letters, 2017

[3] Guillaume Heusch, André Anjos and Sébastien Marcel, A reproducible study on remote heart rate measurement, in: arXiv, 2017

[4] Zeestraten, I. Havoutis, J. Silverio, S. Calinon and D. G. Caldwell, An Approach for Imitation Learning on Riemannian Manifolds, in: IEEE Robotics and Automation Letters (RA-L), 2(3):1240–1247, 2017

[5] Rui Hu, Carlos Pallan Gayol, Jean-Marc Odobez and Daniel Gatica-Perez, Analyzing and Visualizing Ancient Maya Hieroglyphics Using Shape: from Computer Vision to Digital Humanities, in: Digital Scholarship in the Humanities, 32:179–194, 2017

[6] Zahid Akhtar, Abdenour Hadid, Mark Nixon, Massimo Tistarelli, Jean-Luc Dugelay and Sébastien Marcel, Biometrics: In Search of Identity and Security (Q & A), in: IEEE MultiMedia, PP, 2017

[7] Joan-Isaac Biel, Nathalie Martin, David Labbe and Daniel Gatica-Perez, Bites'n'Bits: Inferring Eating Behavior from Contextual Mobile Data, in: Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (PACM IMWUT), 1(4):125–157, 2017

[8] Milos Cernak, Juan Rafael Orozco-Arroyave, Frank Rudzicz, Heidi Christensen, Juan Camilo Vasquez-Correa and Elmar Nöth, Characterisation of voice quality of Parkinson's disease using differential phonological posterior features, in: Computer Speech and Language, 2017

[9] Laurent Son Nguyen, Salvador Ruiz-Correa, Marianne Schmid Mast and Daniel Gatica-Perez, Check Out This Place: Inferring Ambiance from Airbnb Photos, in: IEEE transactions on Multimedia, PP(99), 2017

[10] Milos Cernak, Afsaneh Asaei and Alexandre Hyafil, Cognitive Speech Coding, in: IEEE Signal Processing Magazine, 2017

[11] N. Jaquier, M. Connan, C. Castellini and S. Calinon, Combining Electromyography and Tactile Myography to Improve Hand and Wrist Activity Detection in Prostheses, in: Technologies, 5(4), 2017

[12] Amir Mohammadi, Sushil Bhattacharjee and Sébastien Marcel, Deeply Vulnerable -- a study of the robustness of face recognition to presentation attacks, in: IET (The Institution of Engineering and Technology) -- Biometrics:1--13, 2017

[13] Florian Labhart, Flavio Tarsetti, Olivier Bornet, Darshan Santani, Jasmine Truong, Sara Landolt, Daniel Gatica-Perez and Emmanuel Kuntsche, Development of the Geographical Proportional-to-size Street-Intercept Sampling (GPSIS) method for recruiting urban night-life-goers in an entire city, in: International Journal of Social Research Methodology, 20(6):721–736, 2017

[14] Kevin G. Chan and Michael Liebling, Direct inversion algorithm for focal plane scanning optical projection tomography, in: Biomedical Optics Express, 2017

[15] Dario Azzimonti and David Ginsbourger, Estimating orthant probabilities of high dimensional Gaussian vectors with an application to set estimation, in: Journal of Computational and Graphical Statistics, 2017

[16] Nikolaos Pappas and Andrei Popescu-Belis, Explicit Document Modeling through Weighted Multiple-Instance Learning, in: Journal of Artificial Intelligence Research, 58:591--626, 2017

[17] Rui Hu, Jean-Marc Odobez and Daniel Gatica-Perez, Extracting Maya Glyphs from Degraded Ancient Documents via Image Segmentation, in: Journal on Computing and Cultural Heritage, 10, 2017

[18] Pavel Korshunov and Sébastien Marcel, Impact of score fusion on voice biometrics and presentation attack detection in cross-database evaluations, in: IEEE Journal of Selected Topics in Signal Processing, 11(4):695 – 705, 2017

[19] E. Pignat and S. Calinon, Learning adaptive dressing assistance from human demonstration, in: Robotics and Autonomous Systems, 93:61–75, 2017

[20] D. Bruno, S. Calinon and D. G. Caldwell, Learning Autonomous Behaviours for the Body of a Flexible Surgical Robot, in: Autonomous Robots, 41(2):333–347, 2017

[21] Hannah Muckenhirn, Pavel Korshunov, Mathew Magimai.-Doss and Sébastien Marcel, Long-Term Spectral Statistics for Voice Presentation Attack Detection, in: IEEE/ACM Transactions on Audio, Speech and Language Processing, 25(11):2098–2111, 2017

[22] Riwal Lefort, L. Fusco, O. Pertz and Francois Fleuret, Machine learning-based tools to model and to remove the off-target effect, in: Pattern Analysis and Applications, 20(1):87–100, 2017

Book and book chapters 2017

[23] Nikolaos Pappas, Miriam Redi, Mercan Topkara, Hongyi Liu, Brendan Jou, Tao Chen and Shih-Fu Chang, Multilingual Visual Sentiment Concept Clustering and Analysis, in: International Journal of Multimedia Information Retrieval, 2017

[24] Juan Rafael Orozco-Arroyave, Juan Camilo Vasquez-Correa, Jesús Francisco Vargas-Bonilla, Raman Arora, Najim Dehak, Phani Sankar Nidadavolu, Heidi Christensen, Frank Rudzicz, Maria Yancheva, Alyssa Vann, Nikolai Vogler, Tobias Bocklet, Milos Cernak, Julius Hannink and Elmar Nöth, NeuroSpeech: An open-source software for Parkinson's speech analysis, in: Digital Signal Processing, 2017

[25] Niklas Linde, David Ginsbourger, James Iriving, Fabio Nobile and Arnaud Doucet, On uncertainty quantification in hydrogeology and hydrogeophysics, in: Advances in Water Resources, 110:166–181, 2017

[26] Afsaneh Asaei, Milos Cernak and Hervé Bourlard, Perceptual Information Loss due to Impaired Speech Production, in: IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2017

[27] Aleksandra Cerekovic, Oya Aran and Daniel Gatica-Perez, Rapport with Virtual Agents: What do Human Social Cues and Personality Explain?, in: IEEE Transactions on Affective Computing, 8(3):382–395, 2017

[28] Salvador Ruiz-Correa, Darshan Santani, Beatriz Ramirez Salazar, Itzia Ruiz-Correa, Fatima Rendon-Huerta, Carlo Olmos Carrillo, Brisa Sandoval Mexicano, Angel Arcos Garcia, Rogelio Hasimoto-Beltran and Daniel Gatica-Perez, SenseCityVity: Mobile Crowdsourcing, Urban Awareness, and Collective Action in Mexico, in: IEEE Pervasive Computingg, Special Issue on Smart Cities, 16(2):44–53, 2017

[29] Subhadeep Dey, Petr Motlicek, Srikanth Madikeri and Marc Ferras, Template-matching for Text-dependent Speaker Verification, in: Speech Communication, 2017

[30] R. Reiter-Palmon, T. Sinha, J. Gevers, Jean-Marc Odobez and G. Volpe, Theories and Models of Teams and Group, in: Small Group Research, 48(5):544--567, 2017

[1] Daniel Gatica-Perez, Oya Aran and Dinesh Babu Jayagopi, Analysis of Small Groups, in: Social Signal Processing, pages 349–367, Cambridge University Press. Editors J. Burgoon, N. Magnenat-Thalmann, M. Pantic, and A. Vinciarelli, 2017

[2] Daniel Gatica-Perez, Gulcan Can, Rui Hu, Stephane Marchand-Maillet, Jean-Marc Odobez, Carlos Pallan Gayol and Edgar Roman-Rangel, MAAYA: Multimedia Methods to Support Maya Epigraphic Analysis, in: Arqueologia computacional: Nuevos enfoques para el analisis y la difusion del patrimonio cultural, INAH-RedTDPC, 2017

[3] Pavel Korshunov and Sébastien Marcel, Presentation attack detection in voice biometrics, in: User-Centric Privacy and Security in Biometrics, The Institution of Engineering and Technology, 2017

Articles in conferences proceedings 2017

[1] Trung Phan Thanh and Daniel Gatica-Perez, #Healthy #Fondue #Dinner: Analysis and Inference of Food and Drink Consumption Patterns on Instagram, in: Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia, 2017

[2] Z. Boulkenafet, J. Komulainen, Z. Akhtar, A. Benlamoudi, SE. Bekhouche, F. Dornaika, A. Ouafi, Amir Mohammadi, Sushil Bhattacharjee and Sébastien Marcel, A Competition on Generalized Software-based Face Presentation Attack Detection in Mobile Scenarios, in: Proceedings of the International Joint Conference on Biometrics, 2017, 2017

[3] Youssef Oualil, Dietrich Klakow, Gyorgy Szaszak, Ajay Srinivasamurthy, Hartmut Helmke and Petr Motlicek, A Context-Aware Speech recognition and Understanding System for Air Traffic Control Domain, in: Proceedings of the IEEE Automatic Speech Recognition and Understanding Workshop, Okinawa, Japan, 2017

[4] Nam Le and Jean-Marc Odobez, A Domain Adaptation Approach to Improve Speaker Turn Embedding Using Face Representation, in: ACM International Conference on Multimodal Interaction, Glasgow, Scotland, ACM, 2017

[5] Ajay Kumar Tanwani and S. Calinon, A Generative Model for Intention Recognition and Manipulation Assistance in Teleoperation, in: IEEE/RSJ International Conference on Intelligent Robots and Systems, 2017

[6] James Newling and Francois Fleuret, A Sub-Quadratic Exact Medoid Algorithm, in: Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, 2017

[7] Jagannadan Varadarajan, R Subramanian, Narendra Ahuja, Pierre Moulin and Jean-Marc Odobez, Active Online Anomaly Detection using Dirichlet Process Mixture Model and Gaussian Process Classification, in: IEEE Winter Conference on Applications of Computer Vision (WACV), Washington, 2017

[8] Sibo Tong, Philip N. Garner and Hervé Bourlard, An Investigation of Deep Neural Networks for Multilingual Speech Recognition Training and Adaptation, in: Proc. of Interspeech, 2017

[9] André Anjos, Laurent El Shafey and Sébastien Marcel, BEAT: An Open-Science Web Platform, in: Thirty-fourth International Conference on Machine Learning, Sydney, Australia, 2017

[10] Milos Cernak, Alain Komaty, Amir Mohammadi, André Anjos and Sébastien Marcel, Bob Speaks Kaldi, in: Proc. of Interspeech, 2017

[11] Samaneh Abbasi-Sureshjani, Behdad Dasht Bozorg, Bart ter Haar Romeny and Francois Fleuret, Boosted Exudate Segmentation in Retinal Images using Residual Nets, in: Proceedings of the MICCAI Workshop on Ophthalmic Medical Image Analysis, 2017

[12] Xiao Pu, Laura Mascarell and Andrei Popescu-Belis, Consistent Translation of Repeated Nouns using Syntactic and Semantic Cues, in: Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL), Valencia, Spain, Association for Computational Linguistics, 2017

[13] Subhadeep Dey, Srikanth Madikeri, Petr Motlicek and Marc Ferras, Content Normalization for Text-dependent Speaker Verification, in: Proc. of Interspeech, 2017

[14] André Anjos, Manuel Günther, Tiago de Freitas Pereira, Pavel Korshunov, Amir Mohammadi and Sébastien Marcel, Continuously Reproducing Toolchains in Pattern Recognition and Machine Learning Experiments, in: Thirty-fourth International Conference on Machine Learning, Sidney, Australia, 2017

[15] Ana Sequeira, Lulu Chen, James Ferryman, Peter Wild, Fernando Alonso-Fernandez, Josef Bigün, Kiran B. Raja, R. Raghavendra, Christoph Busch, Tiago de Freitas Pereira, Sébastien Marcel, Sushree Sangeeta Behera, Mahesh Gour and Vivek Kanhangad, Cross-Eyed 2017: Cross-Spectral Iris/Periocular Recognition Competition., in: IEEE/IAPR International Joint Conference on Biometrics, Denver, Colorado, USA, IEEE, 2017

[16] Tatjana Chavdarova and Francois Fleuret, Deep Multi-Camera People Detection, in: Proceedings of the IEEE International Conference on Machine Learning and Applications, 2017

[17] Pierre Baqué, Francois Fleuret and Pascal Fua, Deep Occlusion Reasoning for Multi-Camera Multi-Target Detection, in: Proceedings of the IEEE International Conference on Computer Vision, 2017

[18] D. Berio, S. Calinon and F. F. Leymarie, Dynamic Graffiti Stylisation with Stochastic Optimal Control, in: Intl Workshop on movement and computing (MOCO), London, UK, pages 1–8, ACM, 2017

[19] Rui Hu, Hieu Pham, Philipp Buluschek and Daniel Gatica-Perez, Elderly People Living Alone: Detecting Home Visits with Ambient and Wearable Sensing, in: In Proceedings of MMHealth, 2017

[20] Hannah Muckenhirn, Mathew Magimai.-Doss and Sébastien Marcel, End-to-End Convolutional Neural Network-based Voice Presentation Attack Detection, in: International Joint Conference on Biometrics, Denver, Colorado, USA, 2017

[21] Skanda Muralidhar and Daniel Gatica-Perez, Examining Linguistic Content and Skill Impression Structure for Job Interview Analytics in Hospitality, in: Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia, 2017

[22] Pranay Dighe, Afsaneh Asaei and Hervé Bourlard, Exploiting Eigenposteriors for Semi-supervised Training of DNN Acoustic Models with Sequence Discrimination, in: Proceedings of Interspeech, 2017

[23] Subhadeep Dey, Petr Motlicek, Srikanth Madikeri and Marc Ferras, EXPLOITING SEQUENCE INFORMATION FOR TEXT-DEPENDENT SPEAKER VERIFICATION, in: Proceedings of 2017 IEEE International Conference on Acoustics, Speech, and Signal Processing, New Orleans, pages 5370–5374, 2017

[24] Samaneh Abbasi-Sureshjani, Behdad Dasht Bozorg, Bart ter Haar Romeny and Francois Fleuret, Exploratory Study on Direct Prediction of Diabetes using Deep Residual Networks, in: Proceedings of the thematic conference on computational vision and medical image processing, 2017

[25] N. Jaquier and S. Calinon, Gaussian Mixture Regression on Symmetric Positive Definite Matrices Manifolds: Application to Wrist Motion Estimation with sEMG, in: Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems, 2017

[26] D. Berio, S. Calinon and F. F. Leymarie, Generating Calligraphic Trajectories with Model Predictive Control, in: Proc. 43rd Conf. on Graphics Interface, Edmonton, AL, Canada, pages 132–139, 2017

[27] Skanda Muralidhar, Marianne Schmid Mast and Daniel Gatica-Perez, How May I Help You? Behavior and Impressions in Hospitality Service Encounters, in: Procceddings of 19th ACM International Conference on Multimodal Interaction, 2017

[28] N. Jaquier, C. Castellini and S. Calinon, Improving hand and wrist activity detection using tactile sensors and tensor regression methods on Riemannian manifolds, in: Proc. of the Myoelectric Control Symposium, 2017

[29] Nam Le and Jean-Marc Odobez, Improving speaker turn embedding by crossmodal transfer learning from face embedding, in: ICCV Workshop on Computer Vision for Audio-Visual Media, 2017

[30] Darshan Santani, Salvador Ruiz-Correa and Daniel Gatica-Perez, Insiders and Outsiders: Comparing Urban Impressions between Population Groups, in: International Conference on Multimedia Retrieval, ACM, 2017

[31] Srikanth Madikeri, Marc Ferras, Petr Motlicek and Subhadeep Dey, INTRA-CLASS COVARIANCE ADAPTATION IN PLDA BACK-ENDS FOR SPEAKER VERIFICATION, in: Proceedings of International Conference on Acoustics, Speech and Signal Processing, pages 5365–5369, 2017

[32] James Newling and Francois Fleuret, K-Medoids For K-Means Seeding, in: Proceedings of the international conference on Neural Information Processing Systems, 2017

[33] L. Rozo, N. Jaquier, S. Calinon and D. G. Caldwell, Learning Manipulability Ellipsoids for Task Compatibility in Robot Manipulation, in: Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems, 2017

[34] M. Zeestraten, I. Havoutis, S. Calinon and D. G. Caldwell, Learning Task-Space Synergies using Riemannian Geometry, in: Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS), Vancouver, Canada, pages 73–78, IEEE, 2017

[35] Pranay Dighe, Afsaneh Asaei and Hervé Bourlard, Low-rank and Sparse Soft Targets to Learn Better DNN Acoustic Models, in: Proceedings of 2017 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2017), 2017

[36] Matthias Kleinert, Hartmut Helmke, Gerald Siol, heiko Ehr, Michael Finke, Youssef Oualil and Ajay Srinivasamurthy, Machine Learning of Controller Command Prediction Models from Recorded Radar Data and Controller Speech Utterances, in: Proceedings of the 7th SESAR Innovation Days (SID), University of Belgrade, Belgrade, Serbia, 2017

[37] Ngoc-Quang Luong and Andrei Popescu-Belis, Machine translation of Spanish personal and possessive pro-nouns using anaphora probabilities, in: Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL), Valencia, Spain, Association for Computational Linguistics, 2017

[38] Pierre Baqué, Francois Fleuret and Pascal Fua, Multi-Modal Mean-Fields via Cardinality-Based Clamping, in: Proceedings of the IEEE international conference on Computer Vision and Pattern Recognition, 2017

[39] Juan Camilo Vasquez-Correa, Juan Rafael Orozco-Arroyave, Raman Arora, Elmar Nöth, Najim Dehak, Heidi Christensen, Frank Rudzicz, Tobias Bocklet, Milos Cernak, Hamidreza Chinaei, Julius Hannink, Phani Sankar Nidadavolu, Maria Yancheva, Alyssa Vann and Nikolai Vogler, Multi-view Representation Learning Via GCCA for Multimodal Analysis of Parkinson's Disease, in: Proceedings of 2017 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2017), 2017

[40] Nikolaos Pappas and Andrei Popescu-Belis, Multilingual Hierarchical Attention Networks for Document Classification, in: 8th International Joint Conference on Natural Language Processing (IJCNLP), pages 1015-1025, 2017

[41] Andrii Maksai, Xinchao Wang, Francois Fleuret and Pascal Fua, Non-Markovian Globally Consistent Multi-Object Tracking, in: Proceedings of the IEEE International Conference on Computer Vision, 2017

[42] Sébastien Marmin, Jean Baccou, Jacques Liandrat and David Ginsbourger, Non-parametric warping via local scale estimation for non-stationary Gaussian process modelling, in: Wavelets and Sparsity XVII, pages 1039421, International Society for Optics and Photonics, 2017

[43] Skanda Muralidhar, On Job Training: Automated Interpersonal Behavior Assessment & Real-Time Feedback, in: Proceedings of the 25th ACM International Conference on Multimedia, 2017, 2017

[44] Andreé R. Goncalves, Pavel Korshunov, Ricardo P. V. Violato, Flávio O. Simões and Sébastien Marcel, On the Generalization of Fused Systems in Voice Presentation Attack Detection, in: 16th International Conference of the Biometrics Special Interest Group, 2017

[45] Milos Cernak, Elmar Nöth, Frank Rudzicz, Heidi Christensen, Juan Rafael Orozco-Arroyave, Raman Arora, Tobias Bocklet, Hamidreza Chinaei, Julius Hannink, Phani Sankar Nidadavolu, Juan Camilo Vasquez, Maria Yancheva, Alyssa Vann and Nikolai Vogler, On the Impact of Non-modal Phonation On Phonological Features, in: Proceedings of 2017 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2017), 2017

[46] Sébastien Marmin, Jean Baccou, Frédéric Perales, David Ginsbourger and Jacques Liandrat, Planification adaptative d'expériences numériques par paquets en contexte non stationnaire pour une étude de fissuration mécanique, in: 23eme Congres Francais de Mecanique, 28 aout – 1er septembre 2017, Lille, France (FR), AFM, 2017

[47] Yu Yu, Kenneth Alberto Funes Mora and Jean-Marc Odobez, Robust and Accurate 3D Head Pose Estimation through 3DMM and Online Head Model Reconstruction, in: Proceedings of the 12th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2017), 2017

[48] Ajay Srinivasamurthy, Petr Motlicek, Ivan Himawan, Gyorgy Szaszak, Youssef Oualil and Hartmut Helmke, Semi-supervised Learning with Semantic Knowledge Extraction for Improved Speech Recognition in Air Traffic Control, in: Proceedings of Interspeech 2017, Stockholm, Sweden, pages 2406–2410, 2017

[49] Xiao Pu, Nikolaos Pappas and Andrei Popescu-Belis, Sense-Aware Statistical Machine Translation using Adaptive Context-Dependent Clustering, in: Proceedings of Second Conference on Machine Translation (WMT17), 2017

[50] Gulcan Can, Jean-Marc Odobez and Daniel Gatica-Perez, Shape Representations for Maya Codical Glyphs: Knowledge-driven or Deep?, in: 15th International Workshop on Content-Based Multimedia Indexing, 2017

[51] Timur Bagautdinov, Alexandre Alahi, Francois Fleuret, Pascal Fua and Sylvio Savarese, Social Scene Understanding: End-to-End Multi-Person Action Localization and Collective Activity Recognition, in: Proceedings of the IEEE international conference on Computer Vision and Pattern Recognition, 2017

[52] Afsaneh Asaei, Milos Cernak, Hervé Bourlard and Dhananjay Ram, Sparse Pronunciation Codes for Perceptual Phonetic Information Assessment, in: Workshop on Signal Processing with Adaptive Sparse Structured Representations (SPARS), 2017

[53] Dhananjay Ram, Afsaneh Asaei and Hervé Bourlard, Subspace Regularized Dynamic Time Warping for Spoken Query Detection, in: Workshop on Signal Processing with Adaptive Sparse Structured Representations (SPARS), 2017

[54] I. Havoutis and S. Calinon, Supervisory teleoperation with online learning and optimal control, in: Proc. IEEE Intl Conf. on Robotics and Automation (ICRA), Singapore, pages 1534–1540, IEEE, 2017

[55] Renars Liepins and et al., The SUMMA Platform Prototype, in: Proceedings of the EACL 2017 Software Demonstrations, Valencia, Spain, pages 116--119, 2017

[56] Khaled Khelif, yann Mombrun, Gerhard Backfried, Farhan Sahito, Luca Scarpatto, Petr Motlicek, Damien Kelly, Gideon Hazzani, Emmanouil Chatzigavriil and Srikanth Madikeri, Towards a breakthrough Speaker Identification approach for Law Enforcement Agencies: SIIP, in: European Intelligence and Security Informatics Conference (EISIC) 2017, Athenes, Greece, pages 32–39, IEEE Computer Society, 2017

[57] Nam Le, Jean-Marc Odobez and et al., Towards large scale multimedia indexing: A case study on person discovery in broadcast news, in: 15th International Workshop on Content-Based Multimedia Indexing, 2017

[58] Remy Siegfried, Yu Yu and Jean-Marc Odobez, Towards the Use of Social Interaction Conventions as Prior for Gaze Model Adaptation, in: Proceedings of 19th ACM International Conference on Multimodal Interaction, pages 9, ACM, 2017

[59] C. Mastalli, M. Focchi, I. Havoutis, A. Radulescu, S. Calinon, J. Buchli, D. G. Caldwell and C. Semini, Trajectory and Foothold Optimization using Low-Dimensional Models for Rough Terrain Locomotion, in: Proc. IEEE Intl Conf. on Robotics and Automation (ICRA), Singapore, pages 1096–1103, IEEE, 2017

[60] Lesly Miculicich Werlen and Andrei Popescu-Belis, Using Coreference Links to Improve Spanish-to-English Machine Translation, in: Proceedings of the 2nd Workshop on Coreference Resolution Beyond OntoNotes (CORBON 2017), Valencia, Spain, pages 30-40, Association for Computational Linguistics, 2017

[61] Lesly Miculicich Werlen and Andrei Popescu-Belis, Validation of an Automatic Metric for the Accuracy of Pronoun Translation (APT), in: Proceedings of the Third Workshop on Discourse in Machine Translation (DiscoMT), Denmark, Copenhagen, Association for Computational Linguistics (ACL), 2017

[62] Yassir Benkhedda, Darshan Santani and Daniel Gatica-Perez, Venues in Social Media: Examining Ambiance Perception Through Scene Semantics, in: Proceedings of the 25th ACM International Conference on Multimedia, ACM, 2017, 2017

[63] Stepan Tulyakov, Anton Ivanov and Francois Fleuret, Weakly Supervised Learning of Deep Metrics for Stereo Reconstruction, in: Proceedings of the IEEE International Conference on Computer Vision, 2017

[64] Sushil Bhattacharjee and Sébastien Marcel, What you can't see can help you -- extended-range imaging for 3D-mask presentation attack detection, in: Proceedings of the 16th International Conference on Biometrics Special Interest Group., Darmstadt (Germany), Gesellschaft fuer Informatik e.V. (GI), 2017

PHD THESES 2017

[1] Pierre-Edouard Honnet, Intonation Modelling for Speech Synthesis and Emphasis Preservation, École Polytechnique Fédérale de Lausanne, 2017

[2] Pedro H. O. Pinheiro, Large-Scale Image Segmentation with Convolutional Networks, Sciences et Techniques de l'Ingénieur (STI), 2017

[3] Olivier Canévet, Object Detection with Active Sample Harvesting, École Polytechnique Fédérale de Lausanne, 2017

[4] Marzieh Razavi, On Modeling the Synergy Between Acoustic and Lexical Information for Pronunciation Lexicon Development, École polytechnique fédérale de Lausanne (EPFL), 2017

[5] Gulcan Can, Visual Analysis of Maya Glyphs via Crowdsourcing and Deep Learning, École Polytechnique Fédérale de Lausanne, 2017