Contents

List of Figures

1   Idiap Structure and Research Areas

1.1   Research Areas

Idiap's research activities span five broad areas, all directly related to our central themes and competencies of **Signal Processing**, **Machine Learning**, and **"Social Intelligence for Society"**: perceptual and cognitive systems, human and social behavior, multimodal/multimedia information processing and presentation interfaces, biometrics security and privacy, and machine learning, which are summarized, with a set of descriptive keywords in Figure 1 below.

In addition to these transverse research themes, new activities have been recently developed, towards biomedical applications (through the new group in Computational Bioimaging) and environmental modeling (through the new group in Uncertainty Quantification and Optimal Design).

Idiap maintains a policy of continuous growth and diversification in new and complementary research directions, while building upon its long-standing institutional theme of "Human and Media Computing".

Research Areas	Keywords
Perceptual and cognitive systems	*Speech and audio processing, computer vision, document processing, robotics, natural language processing, machine translation, computational cognitive science*
Human and social behavior	*Social media, verbal and nonverbal communication analysis, smartphone sensing, computational social science*
Information and presentation interfaces	*Multimedia information systems, user interfaces, personalization, system evaluation, mobile HCI using big data, data driven services*
Biometrics Security and Privacy	*Face recognition, speaker recognition, vein recognition, multimodal fusion, soft-biometrics, remore photoplethysmography, presentation attack detection (anti-spoofing), template protection, privacy preservation, mobile and wearable biometrics*
Machine learning	*Statistical and neural network based ML, computational efficiency, online learning, multiple sensor processing, big data*

Figure 1: Overview of Idiap research areas. These research areas are currently covered by eleven research groups, listed in Section 1.3 below, and are designed to be as pro-active as possible in covering the applications areas presented in Figure 2.

1.2 Application Areas

Idiap's application areas are listed in Figure 2, together with keywords and pointers to relevant projects. Idiap contributes to a broad range of applications areas, which reflect concrete, ongoing projects.

We would like to draw attention to the difference between our *research* and our *application* areas. While Idiap's research areas (Figure 1) are quite generic and have a strong transverse nature, the application areas are more targeted and take place along well-defined application axes. Although we do not necessarily address at Idiap all aspects within a given application area, we are often in the best position to provide the necessary enabling Information and Communication Technology components.

Application Areas	*Keywords*
Human-human & human-machine interaction	*Voice and gesture controlled devices and robots, hand-free control, spoken language systems, translation systems, social robotics, user profiling*
Exploitation of rich multimedia archives (audio, video, text)	*Hyper-events, semantic indexing, keyword spotting, object detection and recognition, image bank browsing, audio-video content filtering (summarization and recommendation), broadcast data analysis, scanned document analysis, analysis of cultural heritage media*
Collaborative and creative systems	*Remote meeting assistance, smart meeting room, video-conferencing, multimedia indexing and access, cross-lingual collaboration, interaction analysis, dynamics of negotiation*
Healthcare and bio-engineering	*Smart management of patient data, prosthesis (hearing aids, artificial body parts, voice banking), bio-systems modeling, interfaces for impaired users (speech and other inputs), behavior-based health diagnosis, bio-medical document and data processing, smartphone platforms for health, bio-signals and imaging, microscopy*
Entertainment	*Multilingual gaming, remote family games, togetherness*
Mobile computing	*Signal processing for mobile platforms, mobile social networks, participatory sensing*
Security and risk management	*Biometric security, access control, mobile biometry, multi-sensor fusion, speaker identification, video monitoring of areas/activities, natural risk modeling, intrusion detection, crowd management*
Home automation (domotics)	*Multi-sensor activity analysis, adaptation to users' behavior, efficient use of energy, home safety and security*
Energy	*Energy grids, multiple sensor and smart meter networks, large-scale sensor data integration, modeling of behaviors to anticipate demand, safer, cheaper, and cleaner energy production*
Smart cities	*Ecology, environment management, reduction in pollution, traffic and noise, better use of roads*

Figure 2: Idiap application areas with several examples for each of them.

1.3 Internal Structure

1.3.1 Overview

The main research themes, briefly presented in Section 1.1 above, are currently covered by eleven research groups, presented in the following section. The resulting structure is presented in Figure 3 along with the overall structure of Idiap. The leaders of the groups, who are all PIs of research projects, are explicitly indicated.
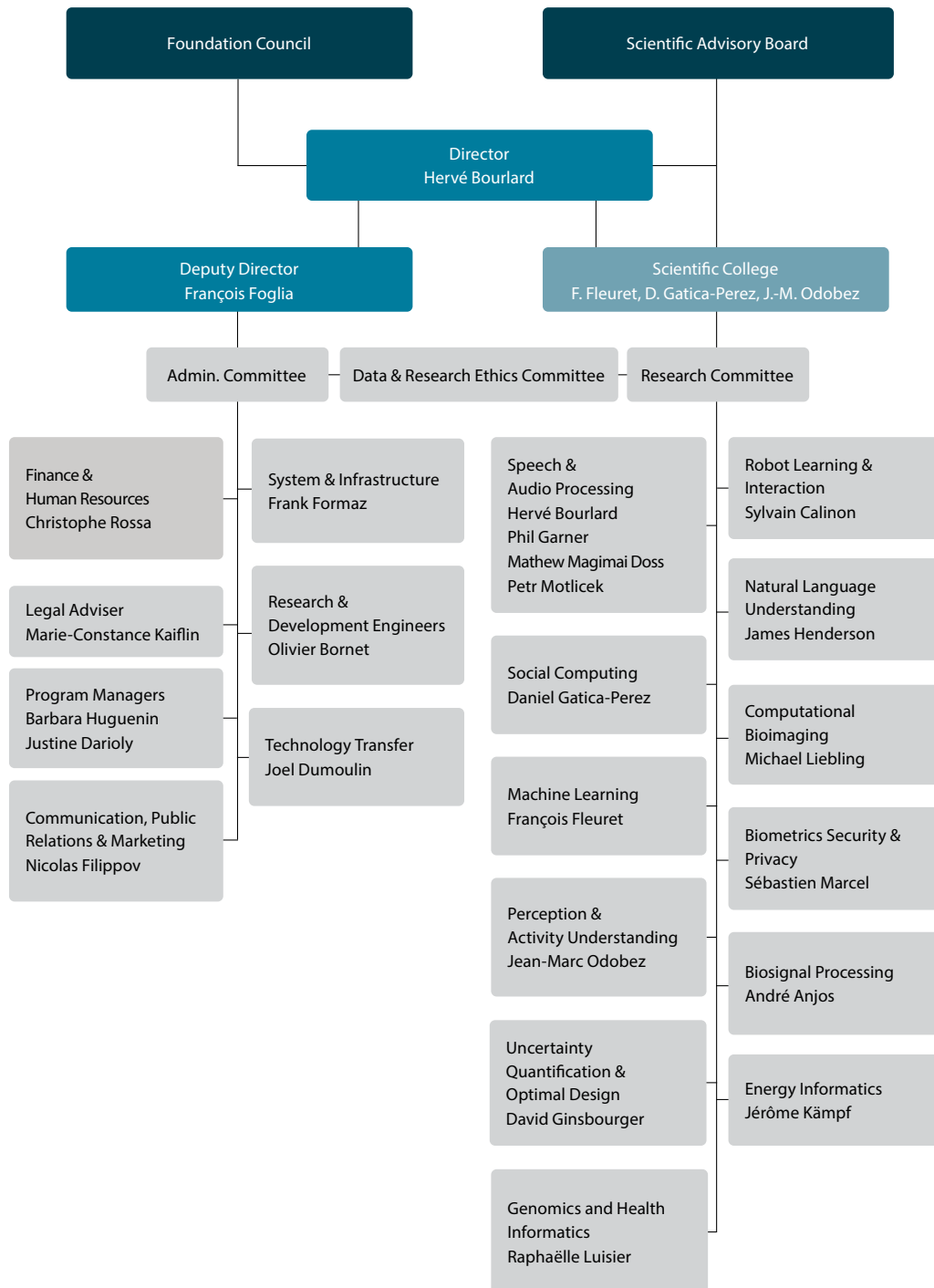

Figure 3: Idiap management and operational structure, including the main research and administrative responsibilities.

1.3.2   Research Groups

In 2019, the research areas presented in Figure 1 were covered by twelve research groups. The activity of each research group (from left to right branch in Figure 1) can be summarized as follows:

1. **Speech and Audio Processing (Prof. Hervé Bourlard, Dr. Philip N. Garner, Dr. Mathew Magimai-Doss, & Dr. Petr Motlicek)**

   Speech processing has been one of the mainstays of Idiap's research portfolio for many years. Today it is still the largest group within the institute, and Idiap continues to be recognized as a leading proponent in the field. The expertise of the group encompasses multilingual speech recognition, multilingual text-to-speech conversion, 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.

   See Section 2.1, page 8, for the 2019 progress report.

2. **Social Computing (Prof. Daniel Gatica-Perez)**

   Social computing is an interdisciplinary domain that integrates theories and models from ubiquitous computing, multimedia, machine learning, and social sciences, to sense, analyze, and interpret human and social behavior in everyday life, and to create systems that support social interaction. Current lines of research include social media analytics and mobile crowdsourcing for cities and health, and analysis of ubiquitous social interaction.

   See Section 2.2, page 14, for the 2019 progress report.

3. **Machine Learning (Prof. François Fleuret)**

   The goal of the Machine Learning group is the development of new statistical learning techniques with a particular interest in their computational properties. Our application domain encompasses any processing of real-world signals, such as image understanding, detection of persons and biological structures, signal synthesis, or decision from low-level industrial sensors.

   See Section 2.3, page 16, for the 2019 progress report.

4. **Perception and Activity Understanding (Dr. Jean-Marc Odobez)**

   The group investigates models from machine learning, computer vision, multimodal signal processing, or social sciences, to address the understanding of activities from real-world signals, with an emphasis on those related to humans. Detection, tracking, pose estimation, recognition and analysis of non-verbal behaviors or the temporal interpretation of all this information in forms of gestures, activities behavior or social relationships are examples of studied tasks. Application domains encompasses surveillance, traffic and human behavior analysis, human-robot interactions, and multimedia content analysis.

   See Section 2.4, page 18, for the 2019 progress report.

5. **Uncertainty Quantification and Optimal Design (Prof. David Ginsbourger)**

   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.

   See Section 2.5, page 20, for the 2019 progress report.

6. **Genomics and Health Informatics Group (Dr. Raphaëlle Luisier)**

The Genomics and Health Informatics Group was created in 2019 to develop statistical and machine-learning methods to integrate genomic, clinical, and imaging data. The group aims to accelerate the diagnosis and improve the understanding and treatment of complex diseases such as neurodegenerative disorders in collaboration with clinical neuroscientists. The group has expertise in genomics, bioinformatics, RNA biology, neuroscience, data science, and data visualisation.

See Section 2.6, page 22, for the 2019 progress report.

7. **Robot Learning and Interaction (Dr. Sylvain Calinon)**

This group focuses on human-centric 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.

See Section 2.7, page 23, for the 2019 progress report.

8. **Natural Language Understanding (Dr. James Henderson)**

The Natural Language Understanding group studies deep learning for natural language processing, focusing on learning representations of the meaning of text and attention-based models of graphs. It models summarization, abstraction (textual entailment), machine translation, knowledge extraction, syntactic structure, and lexical semantics, among other NLP problems.

See Section 2.8, page 25, for the 2019 progress report.

9. **Computational Bioimaging (Prof. Michael Liebling)**

This group focuses on research in computational imaging and 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 for the observation and quantification of complex and live biological systems.

See Section 2.9, page 27, for the 2019 progress report.

10. **Biometrics Security and Privacy (Dr. Sébastien Marcel)**

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

See Section 2.10, page 29, for the 2019 progress report.

11. **Biosignal Processing (Dr. André Anjos)**

Biosignals are signals from living beings and their analysis to support clinical practice and research. This group currently focuses on image (e.g. retinography, X-ray, computerized tomography), vital signs, clinical and multimodal data analysis for healthcare and related applications. Current trends in the field show refreshed interest on the use of machine learning techniques, complementing basic signal and sequence processing, all of which are key domains of research at Idiap. It leverages on Idiap's expertise on human subject handling, data acquisition, open science and data processing.

See Section 2.11, page 33, for the 2019 progress report.

12. **Energy Informatics (Dr. Jérôme Kämpf)**

    The Energy Informatics Group (created September 2018) studies the exploitation of state-of-the-art Information and Communication Technologies to tackle global warming and climate change challenges. The aim is to increase integration of renewable and distributed energy sources by making energy systems smarter, and to increase energy efficiency beyond what improvements at component level can achieve. In that vein, the group researches into ways of simulating energy transition pathways with intelligent control and adjustment mechanisms of evolving buildings with retrofitting and use, renewable energy production and energy storage in a changing climate.

    See Section 2.12, page 35, for the 2019 progress report.

1.3.3   Administration and Services

At the administration and services level, there are seven groups: finances and human resources; communication, public relations and marketing; system and infrastructure; technology transfer; development engineers; program managers; and legal advisor. The resulting admin organization is presented on the left branch of Figure 3, page 3. The main responsibilities of the seven admin groups can be summarized as follows:

1. **Finance and Human Resources (Ed Gregg & Christophe Rossa)**: The Human Resources Department (HR) is integrated within the financial and accounting activities of Idiap and has taken on a greater importance in the past year. With employees from over 30 different countries, the finances and HR department is continually growing to meet the needs of each employee.

2. **Communication, Public Relations & Marketing (Nicolas Filippov)**: The mission of the communication, public relations and marketing department is to use all forms of media and communication to build, maintain, manage the reputation of the Institute, and to promote the Idiap services available for external institutions, such as EU project management, submission proposal tools, etc.

3. **System and Infrastructure (Frank Formaz)**:The main mission of the system and infrastructure group is to provide an optimal and efficient work environment for the Idiap collaborators. The tasks can be split into three main activities covering (1) centralized IT services for the whole Institute (network, storage, servers, workstations, high performance computing, identity management, data distribution), (2) support for collaborators (helpdesk, project specific tasks, web presence), and (3) Infrastructure (building, offices, equipments, central purchasing office).

4. **Technology Transfer and Industrial Relations (Dr Joël Dumoulin)**: Technology transfer is one of the Idiap Research Institute's three core missions. One of the fundamental challenges is to facilitate the interface between the knowledge and the skills of the researcher and the needs of the industrial partners. Idiap resolves this by providing a dedicated multi-disciplinary team of developers and programmers which transfers pieces of software, algorithms, knowledge and expertise. This transfer of technology is usually done by granting rights on the commercial exploitation of this technology (through license). Maintaining the Technology Portfolio and helping the researchers in the process of patenting are among the other activities of the Technology Transfer Officer (TTO). Also in charge of the industrial relations, the TTO is the contact person for companies interested in collaborating with Idiap and presents Idiap and Idiap's activities to them. His role is to identify collaboration opportunities with potentially new partners, initiate the discussion and involve the related researcher at the right time in order to improve the efficiency of the process of preparing industrial projects.

5. **Program Managers (Justine Darioly & Barbara Huguenin)**: The work of the program management team is divided into two types of activities. The first is the provision of services to researchers within the framework of European and Swiss projects. The second category includes activities ranging from event organization to database management, which are not directly linked to the management of research projects but facilitate the work of Idiap researchers.

6. **Development Engineers (Olivier Bornet)**: The mission of the development team is to provide support to Idiap researchers in the software development tasks. This is done in three areas. The first is to help on Idiap research by building prototypes, implement algorithms, design and run experiments, and manage legacy code. The second area is for all the technology transfer tasks. In the third activity, development engineers give daily support to Idiap researchers (software disclosures, showroom and internal demonstrators, development tools).

7. **Legal Adviser (Marie-Constance Kaiflin)**: The main missions of the legal adviser are to write, analyse, and negotiate project contracts (research, consortium, consultant agreements, NDA, Memorandum of Understanding, etc) or technology transfer contracts (patents, knowhow, licenses) with industries, universities or research institutions. The legal adviser deals also with all the legal aspects related to human resources (work contracts, staff regulations, rules) and data protection (ethics, databases collection and distribution).

2   Research Groups

2.1   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 2019 was composed of 1 head of group, 3 principal investigators, 1 sabbatical academic visitor, 2 research associates, 9 postdocs, 12 PhD students, and 7 interns.

Key scientific outputs: Our primary research directions have traditionally been Hidden Markov Modeling (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 Automatic Speech Recognition (ASR) 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 ASR are usually researched through *Kaldi* toolkit, now used by most of the international speech community, or its combination with other deep learning tools (particularly *Pytorch*).

In 2019, several key research contributions were achieved by the group, including: (1) multilingual automatic speech recognition, especially in cross-lingual adaptation, and automatic speech recognition in low-resourced language conditions, (2) speaker recognition, through both text-independent and particularly text-dependent (i.e. particularly for 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, (4) new Compressive Sensing (CS) and Sparse Recovering theories to automatic speech recognition, (5) detection of impairments in speech signal to uncover motor speech disorders, and (5) paralinguistic speech processing with minimal prior knowledge. Beside that, the group is also involved in deployment of speech and speaker recognition algorithms for industrial applications (e.g. keyword spotting detection and speaker identification for embedded devices, automatic speech recognition jointly trained with natural language modeling task, etc.).

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-/multi-lingual processing. More specifically, in addition to focusing on "majority" languages other than English such as, French, German, Idiap is actively carrying research in several ASR directions, including:

- **Robust parametrisation of speech recognition models**: We are 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). In the context of recently completed EC H2020 project MALORCA project[1] and recently funded EC H2020 CleanSky project ATCO2,[2] we are developing semi-supervised learning methods for rapid adaptation of speech recognition models to new (unseen) domains using unlabelled data. We are also developing methods to perform ASR system failure analysis with incomplete data (e.g., lack of transcription) and through use of information-theoretic measures.
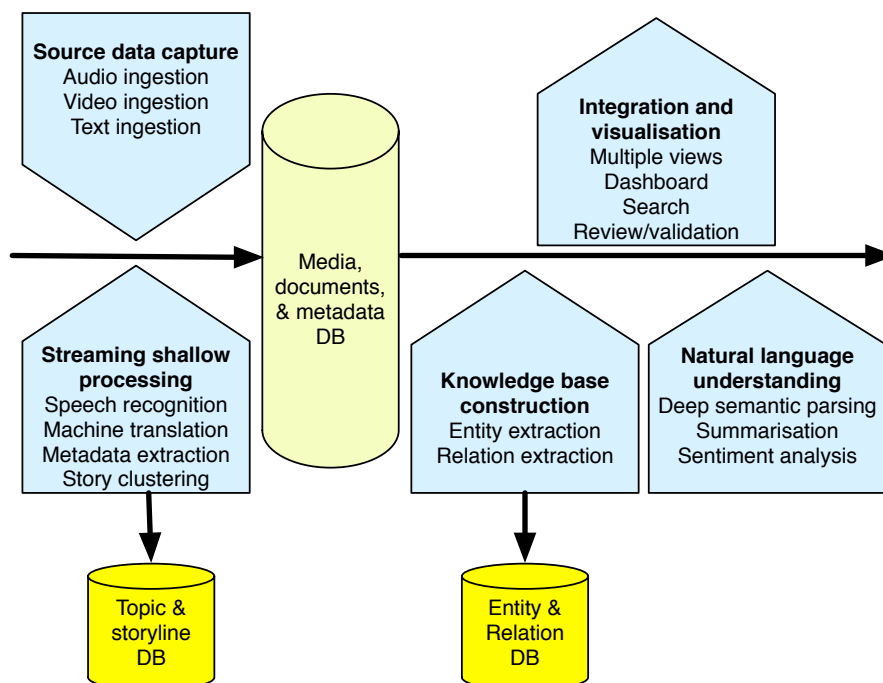
Figure 4: An overview of a typical multilingual multimedia processing stream, as used in EU-H2020 SUMMA and US-DARPA SARAL projects in the context of multilingual speech recognition at Idiap.

- **Cross-lingual and multi-lingual speech recognition (specifically for low-resource scenarios)**: From 2017, Idiap collaborates on the US IARPA SARAL project (Summarization and domain-Adaptive Retrieval of Information Across Languages). As illustrated in Figure 4[3], the project 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 HMM/DNN approach, where Deep Neural Network (DNN) acoustic models are bootstrapped using well-resourced data and adapted to the target language. Recently finished EC SUMMA project (Scalable Understanding of Multilingual Media), aimed at in-

---

[1] http://www.malorca-project.de
[2] http://www.atco2.org
[3] https://www.idiap.ch/en/scientific-research/projects/SARAL

tegrating 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 Deutsche Welle.

- **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. 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 German (dialect) speech recognition. Idiap works on an ongoing Innosuise project SM2[4] ("Extracting semantic meaning from spoken documents") fostering collaboration of Idiap with *recapp IT AG* and *Swiss Learning Hub AG*, with aim to develop a customisable technology for "automatic speech recognition" followed by "semantic keyword and concept detection and spoken document summarization" applied to e-learning domain

- **Exploiting compressive sensing and sparse recovering theories for ASR**: Idiap is still continuing development of new theoretical links between compressive sensing and statistical/HMM-DNN approaches. These efforts are focusing on: (a) improving ASR performance through a link between statistical speech recognition formalism and the recent theories of sparse modelling and (probabilistic) principle component analysis (PCA) and (b) developing new spoken term query detection algorithms that rely on the characterization and detection of the low-dimensional subspace of the DNN phonetic posteriors, through SNSF funded projects PHASER, PHASER-QUAD and SHISSM.[5]

- **Punctuation prediction**: As our ASR output becomes an input for processing at higher semantic levels, for instance in the EU-H2020 SUMMA and US-DARPA SARAL projects, 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.

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, D-BOX and SP2 (SCOPES project on speech prosody). The group has tracked the recent developments in deep learning which will dominate future research. Current work under MASS (Multilingual Affective Speech Synthesis) brings the concept of emotion into the speech synthesis, particularly via modelling of prosody. The research focusses on how to integrate emotional indicators such as pitch, energy and formant position into state of the art deep learning solutions. A new open-source package IdiapTTS[6] was released in 2019 reflecting our work in this area.

- **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

---

[4]https://www.idiap.ch/en/scientific-research/projects/SM2
[5]https://www.idiap.ch/en/scientific-research/projects/SHISSM
[6]https://github.com/idiap/IdiapTTS

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.

Speaker recognition and speech data mining

Idiap is actively carrying R&D on (a) significantly improving capabilities of voice technologies in suspect identification applicable to very large scale data. In that regard, it is worth mentioning that the technologies developed under recently completed EC FP7 SIIP project has successfully passed three proof-of-concept and field-test events, (b) development of methods to extract relevant evidences from speech and audio signals for forensic agencies that may be admissible in a court of law and (c) improving the state-of-the-art speaker technologies by integrating recent advances in machine learning (especially thanks to the participation on NIST speaker recognition evaluations[7]). In late 2019, as a successor to the SIIP project, new EC H2020 security project ROXANNE[8] (funded under the "Fight against crime and Terrorism" programme), coordinated by Idiap, has started with an overall budget of 7M EUR. The ROXANNE project aims at combining the strengths of speaker data mining and network analysis algorithms, further enriched by technologies such as natural language understanding, or video analysis to detect not only individual suspects but also the whole criminal or terrorist networks. Furthermore, as discussed in Section 2.10, and as part of the SNSF Project UniTS,[9] the group is also contributing to the Biometrics Security and Privacy group, including the development of countermeasures to detect attacks on speaker verification systems through forged speech samples.
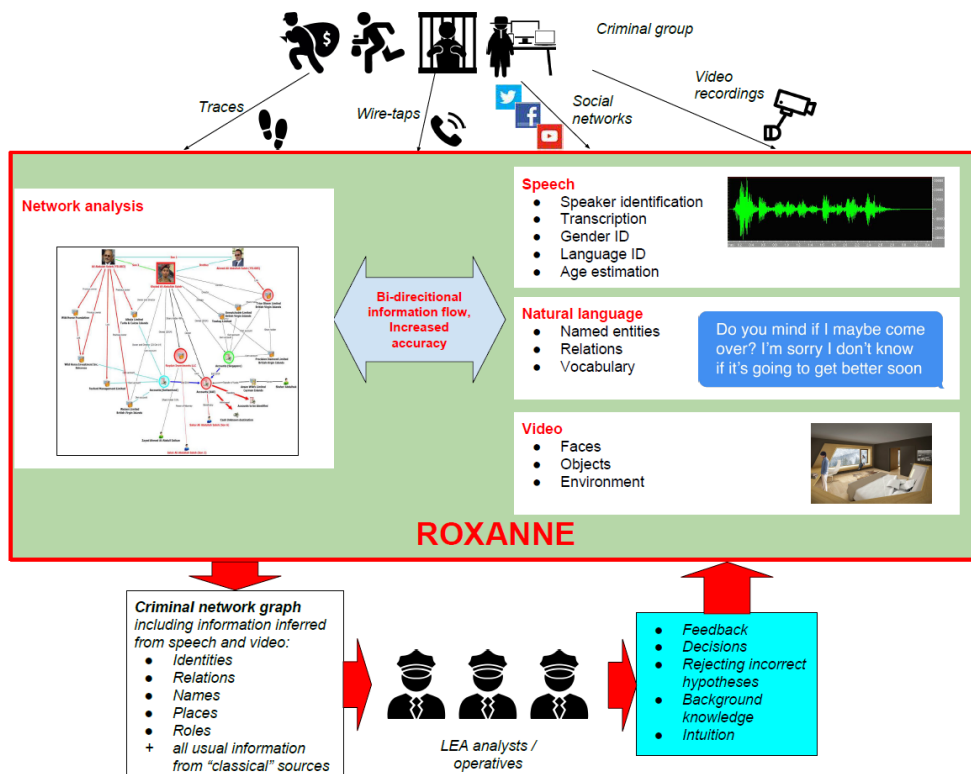

Figure 5: Illustration of ROXANNE platform combining speech data mining with network analysis.

Pathological speech processing

Speech and language impairments can occur due to various reasons such as, due to neurological disorders, oral cancer, hearing loss. In recent years, Idiap has been actively involved in such impaired or pathological speech processing in collaboration with clinical researchers through,

- SNSF Sinergia project MoSpeeDi[10] (Motor Speech Disorder), which focuses on (i) developing accurate models to characterize non-impaired and impaired phonetic speech planning and motor speech programming, and ii) developing automatic techniques to detect and classify several speech impairments as well as to assess the intelligibility of patients, 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. Idiap is addressing the second objective by developing methods for automatic pathological speech detection and objective speech intelligibility assessment.

- EC H2020 MSCA-ITN-ETN project TAPAS[11] (Training Network on Automatic Processing of PAthological Speech), which is 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 (re-designing current speech technology so that it works well for people with speech impairments and also helps in making informed clinical choices. Idiap is contributing towards (a) development of automatic pathological speech analysis and detection and (b) improving impaired speech recognition.

Other directions

- **Sign language processing**: In the context of SNSF Sinergia project SMILE,[12] Idiap, in collaboration with HfH, Zurich and University of Surrey, UK, is developing 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. See, `vimeo.com/297803984`, for a demonstration.

- **Sound localization and microphone array**: Idiap continues to work on distant speech processing by contributing to the Perception and Activity Understanding group through EC H2020 MuMMER project, focusing on audio source localization, speech detection and speaker re-identification applied in robotics.

- **Effective processing of speech using embedded devices**: Integrating voice technologies in low powered devices with limited computing capabilities, Idiap is collaborating with *Logitech* through CTI project SHAPED[13], as well as with *CSEM* to develop a speech hybrid analytics platform for dedicated consumer and enterprise devices or for the healthcare market.

- **Joint acquisition and modeling of speech and physiological signals**: Since late 2018, through internally funded project REAPPS[14] and recently funded SNSF project TIPS[15], Idiap is collaborating with CSEM (`www.csem.ch`) to develop a platform where speech and physiological signals are collected in a synchronous manner through a wearable cooperative sensor and processed to develop novel speech- and physiology-based applications.

---

[10] `https://www.idiap.ch/en/scientific-research/projects/MOSPEEDI`
[11] `https://www.tapas-etn-eu.org`
[12] `https://www.idiap.ch/scientific-research/projects/smile`
[13] `https://www.idiap.ch/en/scientific-research/projects/SHAPED`
[14] `https://www.idiap.ch/en/scientific-research/projects/REAPPS`
[15] `https://www.idiap.ch/en/scientific-research/projects/TIPS`

Key publications

[1] D. Palaz, M. Magimai.-Doss, and R. Collobert, "End-to-end acoustic modeling using convolutional neural networks for HMM-based automatic speech recognition," in *Speech Communication*, vol. 108, pp. 15–32, Apr. 2019

[2] F. Marelli, B. Schnell, H. Bourlard, et al., "An end-to-end network to synthesize intonation using a generalized command response model," in *IEEE International Conference on Acoustics, Speech and Signal Processing* (ICASSP 2019), Brighton, United Kingdom: IEEE, May 2019, pp. 7040–7044

[3] W. He, P. Motlicek, and J.-M. Odobez, "Adaptation of multiple sound source localization neural networks with weak supervision and domain-adversarial training," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing* (ICASSP 2019), Brighton, United Kingdom, May 2019, pp. 770–774

[4] I. Kodrasi and H. Bourlard, "Super-gaussianity of speech spectral coefficients as a potential biomarker for dysarthric speech detection," in *IEEE International Conference on Acoustics, Speech and Signal Processing* (ICASSP 2019), Brighton, United Kingdom: IEEE, May 2019, pp. 7040–7044

[5] S. Tong, A. Vyas, P. N. Garner, et al., "Unbiased semi-supervised LF-MMI training using dropout," in *Proceedings of Interspeech*, 2019

2.2   Social Computing

Overview

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

Group overview: 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 2019 was composed of one group head, three postdoctoral researchers, three PhD students, one scientific collaborator, and two EPFL master students. The main research lines investigated in 2019 included social media analytics and mobile crowdsourcing for cities and health, and ubiquitous conversational interaction analysis.

Key scientific outputs: Publications on (1) social media analytics and mobile and online crowdsourcing to understand youth nightlife and urban perception in cities; (2) multimodal analysis of social interaction. 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

First, in the context of the Dusk2Dawn project[16] (Characterizing Youth Nightlife Spaces, Activities, and Drinks, in collaboration with La Trobe University and the University of Zurich) supported by the Swiss National Science Foundation (SNSF), we investigated the use of mobile crowdsourcing and social media analytics to characterize urban phenomena related to nightlife (Figure 6). This included the automatic recognition of night drinking activity from smartphone sensor data (location, motion, bluetooth, wiki, and app logs) and social media in Switzerland, including a comparison between data sources [1]. Additionally, we studied the phenomenon of place ambiance perception in personal spaces and the automatic recognition of place ambiance from crowdsourced videos using deep learning [2] This research line was complemented, in the context of place-centric online platforms like Airbnb, with qualitative methods used to understand how users relate to sharing their personal spaces [3].

Regarding mobile crowdsourcing and health, as an outcome of the Bites-n-Bits project[17] (in collaboration with Nestle Research Center), we analyzed smartphone sensing data to capture everyday life eating and drinking patterns and learn eating routines of college students [4]. This research is extended in the context of the European H2020 WeNet project[18], which is building diversity-aware algorithms for mobile sensing and social interaction to support the well-being of young people, and is developing a series of large-scale experiments in several European universities, as well as universities in Latin America and Asia. One key motivations of this work is the advocacy for diversity in data and algorithms to improve the representation of non-western citizens[19].

Regarding mobile crowdsourcing for social innovation, we continued our work using the Civique platform[20] (Mobile Data for Local Causes), which has shown its versatility of use, ranging from supporting cities to collect information related to urban issues like street harassment, to teaching students about humanitarian technologies. In 2019, the platform was used in the context of the LOIS project (Leveraging On-Device

---

[16]http://www.idiap.ch/project/dusk2dawn
[17]https://www.bitesnbits.org/
[18]https://www.internetofus.eu/
[19]https://www.lavanguardia.com/lacontra/20190315/461031608874/es-mas-facil-saber-por-que-amamos-que-html
[20]https://www.civique.org
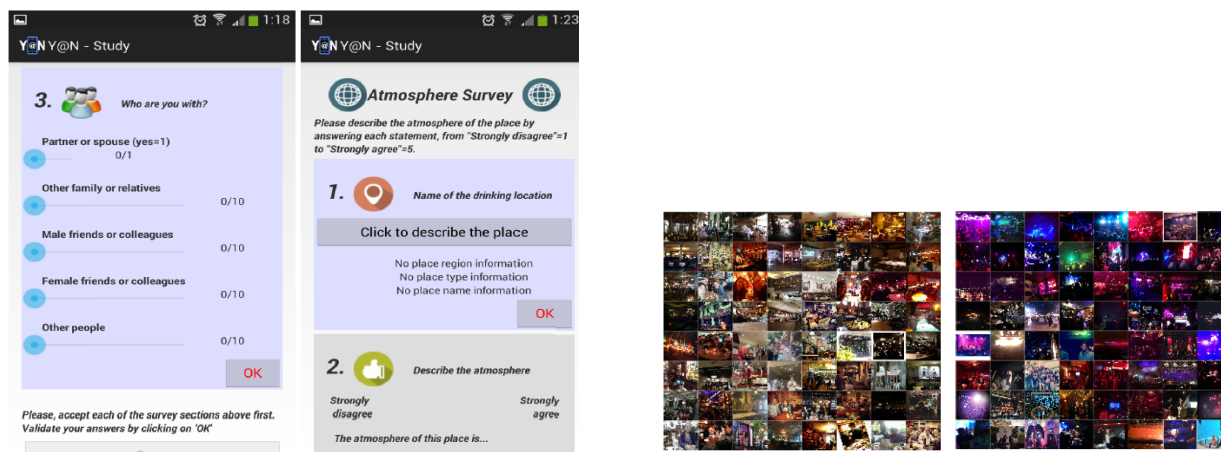
Figure 6: Left: Youth@Night survey logger app. Right: Social media images classified as Restaurant and Stage by convolutional neural networks.

Smartphone Inference for Social Surveys, in collaboration with the University of Lausanne), funded by the UNIL-EPFL Program on Cross-Disciplinary Research on Science and Society. This project explored the potential of smartphones to complement and potentially improve the collection of large-scale, statistically representative survey data by social scientists[21].

Ubiquitous interaction analytics

In the context of the SNSF Digital Lives project[22] (in collaboration with the University of Neuchatel and the University of Lausanne), we are investigating the question of how to provide effective feedback in online video interviews, through the integration of methods that analyze this form of dyadic interaction from audio-visual sensor data. The project was inserted within the national interest to support research addressing questions on the digital transformation of Switzerland[23]. This line of work complements other lines of research investigated in our group regarding other social interaction scenarios, and has resulted in two innovation projects with start-ups supported by Innosuisse.

Key publications

[1] T. T. Phan, S. Muralidhar, and D. Gatica-Perez, "Drinks and Crowds: Characterizing Alcohol Consumption through Crowdsensing and Social Media," in *PACM on Interactive, Mobile, Wearable, and Ubiquitous Technologies* (IMWUT), Vol. 3, No. 2, Jun. 2019

[2] T. T. Phan, F. Labhart, and D. Gatica-Perez, "My Own Private Nightlife: Understanding Youth Personal Spaces from Crowdsourced Video," in *PACM on Human-Computer Interaction* (PACMHCI), Vol. 3, Issue CSCW, Nov. 2019.

[3] S. Ruiz-Correa, I. Ruiz-Correa, C. Olmos, Alba Rendon, B. Ramirez, L. S. Nguyen, and D. Gatica-Perez, "Examining Airbnb Hospitality Exchange Practices in a Developing Economy," in *ACM Trans. Social Computing*, Vol. 2, No. 1, Feb. 2019.

[4] D. Gatica-Perez, J.-I. Biel, D. Labbe, and N. Martin, "Discovering Eating Routines in Context with a Smartphone App," in *Proc. ACM Ubicomp/ISWC'19 Adjunct Proceedings*, Workshop on Mining and Learning from Smartphone Apps, London, Sep. 2019

---

[21]https://actu.epfl.ch/news/cross-researchers-present-their-projects-on-resist/
[22]http://wp.unil.ch/digital-lives/
[23]https://pages.rts.ch/la-1ere/programmes/cqfd/14-03-2019#10253438

2.3 Machine Learning

Overview

Head: Prof. François Fleuret (MS École Normale Supérieure de Paris and University of Paris VI, 1995; PhD, University of Paris VI, France, 2000; EPFL Adjunct Professor)

Group overview: Machine learning encompasses computer techniques that modulate their behavior according to exemplar data. It has resulted in technologies at the core of many modern every-day data-processing software and apparatus. The objective of the Machine Learning group is to develop novel machine-learning techniques of general use, with a particular interest in algorithmic efficiency and training from small data-sets. The research we conduct can be motivated by a general and fundamental problem, or by a concrete industrial application or use case.

Over the recent years, the group has been composed on average of four PhD students, and one or two developers working on industrial applications. We also maintained a sustained collaboration with EPFL's eSpace Center, CVLab and MLO lab.

Key scientific outputs: In 2019, our work has resulted in contributions that improved surface reconstruction from large signal and event-based sensors, interpretability of deep architectures neural networks, more stable training procedures for generative adversarial models, and importance-sampling methods for dealing with very large images.

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

Jacobian matching and regularization

Transfer learning aims at taking advantage of pre-existing models to facilitate the training of new models, either by speeding it up, or by allowing it with very small amount of training data. The key notion is to "transfer" structures learned by the existing network. Our algorithm consists of a novel penalty that not only forces the new model to mimic the response of the existing one, but to also mimic the dynamic of change of the output, given changes of the input. We have put the same tools to use for interpretability, relying on the amplitude of the derivatives of a neural network's output with respect to internal activation to provide a consistent estimate of the importance of different parts of the signal in driving the network's decision.

Stable Adversarial Optimization

"Generative Adversarial Networks" rely on training jointly two models, one synthesizing realistic signals (images, sound, text) and another trying to discriminate synthetic from genuine examples. 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.

Re-sampling for deep models

The computational requirement for deep neural networks is one of their most problematic characteristics, even though most of the computation is spent on samples that are properly handled, and could be ignored. For training, we have derived a tractable upper bound of the per-sample gradient norm that allow to prioritize re-sampling examples and reduce the variance of the stochastic gradient estimates. During inference, we have designed a new model to handle megapixel images, composed of a first network that computes
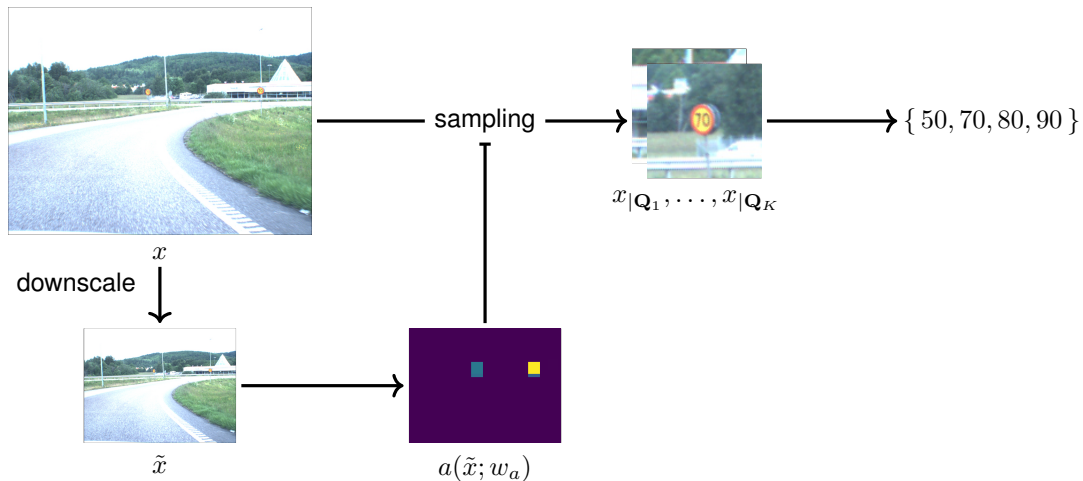
Figure 7: The algorithm we have developed is composed of two models: One processes a downscaled version $\tilde{x}$ of the image in a regular manner to compute an attention map, the second processes only locations $\mathbf{Q}_1, \ldots, \mathbf{Q}_K$ sampled according to the attention of the first. The two networks are optimized end-to-end jointly. On this data-set of speed limit signs, the first network "learned" to look at signs even though this information was never provided during training, and the second network learned to recognize the different speed numbers to make the final prediction.

an attention map on a downscaled input, and a second network that processes the full resolution input at locations sampled according to the attention scores.

Depth estimation with deep models

End-to-end deep-learning networks are now the most efficient method for stereo matching. However, existing networks are memory-hungry and unable to process even modest-size images, and they have to be trained for a specific disparity range. The Deep Stereo network that we developed addresses both issues: First, its architecture relies on novel bottleneck modules that drastically reduce the memory footprint in inference, and additional design choices allow to handle greater image size during training. We have extended this approach to a new class of event-based sensors which provides extremely reactive visual information, at the cost of the point-wise intensity estimation and spatial resolution. Our algorithm relies on a new integration module that combines evidences through time to allow a classical deep architecture to be put to use.

Key publications

[1] T. Chavdarova, G. Gidel, F. Fleuret, and S. Lacoste-Julien, "Reducing Noise in GAN Training with Variance Reduced Extragradient", in *Proceedings of the international conference on Neural Information Processing Systems (NeurIPS)*, pages 391–401, 2019.

[2] S. Srinivas and F. Fleuret, "Full-Gradient Representation for Neural Network Visualization," in *Proceedings of the international conference on Neural Information Processing Systems* (NeurIPS), pages 4126–4135, 2019.

[3] S. Tulyakov, F. Fleuret, M. Kiefel, P. Gehler, and M. Hirsch, "Learning an event sequence embedding for event-based deep stereo," in *Proceedings of the IEEE International Conference on Computer Vision* (ICCV), pages 1527–1537, 2019.

[4] A. Katharopoulos and F. Fleuret, "Processing Megapixel Images with Deep Attention-Sampling Models," in *Proceedings of the International Conference on Machine Learning* (ICML), pages 3282–3291, 2019.

2.4   Perception and Activity Understanding

Overview

Head: Dr. Jean-Marc Odobez (PhD, INRIA/Rennes University, France, 1994; EPFL MER)

Group overview: The group conducts research in human activity 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.

Over the last 10 years, the group was composed on average of two post-doctoral researchers, five PhD students, one research engineer from the development team.

Key scientific outputs: The group is known for its work on probabilistic single and multi-object tracking, non-verbal behavior extraction, and temporal motif discovery. In 2015 and 2016, the PAU team ranked first at the MediaEval Person discovery challenge, related to its investigation on multi-modal person face diarization within the EU EUMSSI project. Its continuing and patended work on 3D face and gaze tracking from cheap RGB-Depth sensors has led to the creation of the Eyeware SA start-up company. Also, the team ranked second at the ICCV 2019 Facebook Synthetic Eye Generation Challenge. In recent years, the group has been investigating 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, and multimedia processing (cross-modal transfer learning, shape recognition, text localization and semantic extraction). It is also working on the integration of its sensing technology and algorithms into real-time perceptual systems used in collaboration projects (human-robot interaction Pepper platform, EU MuMMER project, see video), or by companies (e.g. anti-tailgating detection system). During the period 2015-2019, the group published 15 journal papers, above 35 conference papers, and filled 3 patents.

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

Human activity analysis, non-nerbal behavior (NVB) extraction, human-robot interactions

Head and body pose inference from RGB-D (color and depth) data. A recent achievement has been the development of a robust and accurate head pose tracking framework 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 (Figure 8b), making head tracking a commodity for situations up to 1.5m. Research on 2D and 3D body pose estimation and tracking is also conducted (see Fig. 8). We successfully investigated lightweight and efficient CNN structures as building blocks within a recursive landmark and limb estimation architecture, training with synthetic data and domain adaptation techniques, and knowledge distillation at several architecture levels to increase our lightweight model accuracy [1].
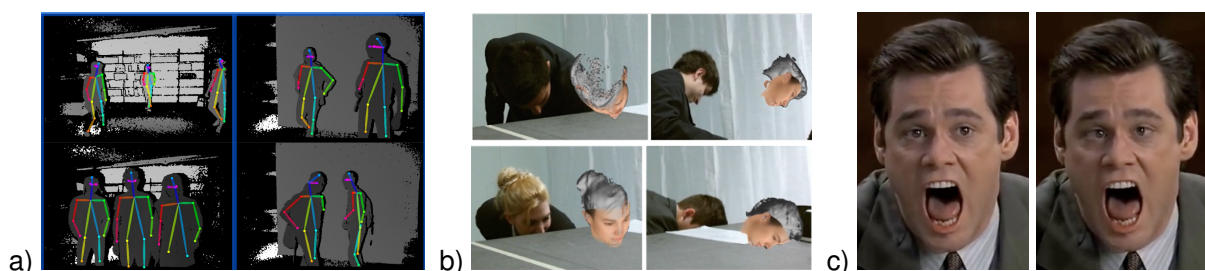

Figure 8: Human activity analysis. a) real-time body landmark detection from depth images. b) accurate and robust 3D head and gaze tracking in adverse conditions. c) example of redirected gaze (left image: original; right one: redirected gaze towards the nose).
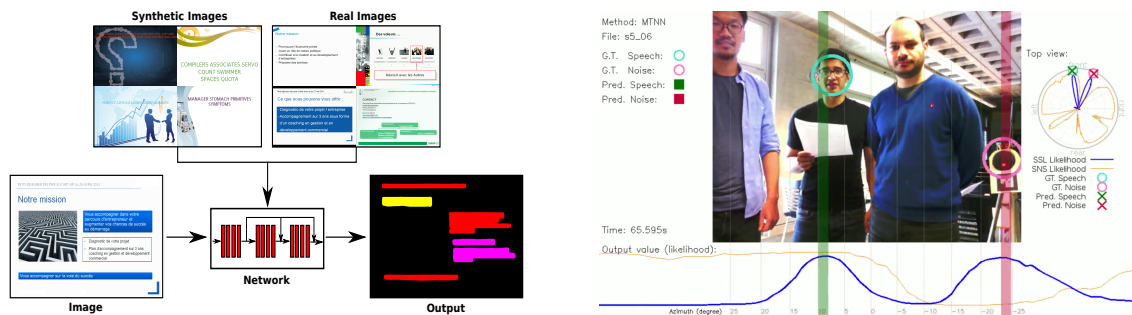
Figure 9: Left: Semantic Text Segmentation Network (STSN), that apart from localizing text in images with high accuracy also classifies text regions in different semantic categories (title, bullets, standard text), and then further recognizes text (OCR). Right: Multiple sound source localization and discrimination for HRI. A multi-task network learns to predict for each direction (circle on the right) the likelihood that there is a sound source (blue curve) and whether it is a speech sound (yellow curve). Sound sources are then mapped to the visual data.

Gaze analytics. Due to visually unobservable variabilities across people regarding the human gaze system, we have continued our work on the building of user-specific models from some user samples. First, we followed the idea of using pairs of eye images of the same person to predict gaze differences (differential gaze). Using two eyes as input better handles nuisance parameters like eye shape or illumination conditions which can impact gaze prediction, and works much better than state-of-the-art methods [2]. In [3], we explored another idea consisting of synthesizing additional user-specific training samples from few calibration ones, by designing an effective gaze redirection network (see illustration in Fig. 8c), and showed that the method could further improve our previous results. Other works include the design of an efficient image-based eye activity segmentation (into blinks, saccades, fixations) which can serve several purposes (temporal gaze filtering, infer the user cognitive state), and online adaptation techniques leveraging gaze priors in several context (social situations, object manipulation tasks) for obtaining weakly labeled gaze samples.

Multimedia and multimodal analysis

Semantic text recognition (OCR). In the Innoswiss VIEW project, we investigated deep learning methods for the detection, segmentation, categorization and recognition of text content in slides, allowing further semantic tagging [4] (see Figure 9). Models and systems have been transfered to the Klewel company, improving the internet referencing and visibiliy of the recorded events of its client.

Audio analysis. Within the EU MuMMER project on social robotics, we are investigating different DNN architectures for sound processing. We proposed an efficient multi-task approach for the joint localization and categorization (speech vs non-speech) of multiple sources from a microphone array, a frequent situation for robots placed in public spaces. It is illustrated in Fig. 9, and was nominated for the best student paper award at Interspeech. More recently, we studied methods to drastically reduce the amount of annotated data needed to train our architectures for each new microphone array [5]: use of simulated data; domain adaptation (from simulation to real data); and a novel weakly supervised scheme (knowing the number of sources, not their location) demonstrating that similar performance to the full supervised case could be obtained with much less (real) training data.

Key publications

[1] A. Martínez-Gonzàlez, M. Villamizar, O. Canévet, J.-M. Odobez, "Efficient Convolutional Neural Networks for Depth-Based Multi-Person Pose Estimation," in *IEEE Trans. Circ. and Sys. for Video Technology*, 2019.

[2] G. Liu, Y. Yu, K. Funes and J.-M Odobez. "A Differential Approach for Gaze Estimation," in *IEEE Trans. Pattern Anal. Mach. Intell.* (accepted for publication), Nov. 2019.

[3] Y. Yu, G. Liu and J.-M Odobez. "Improving Few-Shot User-Specific Gaze Adaptation via Gaze Redirection Synthesis," in *Int Conf. on Vision and Pattern Recognition (CVPR)*, June 2019.

[4] M. Villamizar, O. Canévet, J.-M. Odobez, "Multi-scale sequential network for semantic text segmentation and localization," *Pattern Recognition Letters*, Vol. 129, pp 63–69, 2020.

2.5 Uncertainty Quantification and Optimal Design

Overview

Head: Prof. David Ginsbourger (Ph.D. Mines Saint-Etienne 2009, Habilitation Univ. Bern 2014, Titularprofessor Univ. 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 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 2019, the UQOD group has been composed of a permanent senior researcher, two PhD students, an intern, and three visiting students (one PhD student from the University of Neuchâtel and two UniBE master students).

Key scientific outputs: Current contributions include efficient algorithms for Bayesian optimization and for estimating and quantifying uncertainties on implicitly defined parameter regions using Gaussian Process (GP) models. Other recent results deal with the interplay between the choice of covariance kernels and properties of GPs, with implications in high-dimensional GP modelling, in function prediction under structural constraints, and in problems with set-valued inputs. Ongoing work encompasses uncertainty quantification and efficient design of experiments for data- and simulation-driven inverse problem solving with applications in computational cosmology and geophysics (Current SNSF project).

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) stemming from heavy numerical simulations. More generally, sequential design of (computer) experiments based on GPs have flourished to efficiently address a variety of goals. This 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, with set-valued inputs, etc.) and applications. Notably, the group has been involved in a collaboration with researchers in hydrogeology with the aim to investigate Bayesian optimization for contaminant source localization relying on flow simulations.

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 directions include investigations on positive definite kernels over sets of finite sets using RKHS embeddings, with application to Bayesian (combinatorial) optimization [2]. Fig. 10 represents two instances of a score discrepancy landscape corresponding to an optimal and to an arbitrary subset of five monitoring wells among twenty-five, respectively, in the framework of the aforementioned contaminant localization problem. It is found that Bayesian Optimization algorithms based on the considered RKHS embeddings most of the time successfully locate the optimal well combination (out of a total of 53 130 candidates) in a few dozens of iterations. Furthermore, UQOD has been involved in collaborations around the adaptation of GP models and Bayesian Optimization algorithms to challenging frameworks, be it for instance in terms of high-dimensionality or of heteroscedastic multivariate trajectories in the context of robot learning. The former topic has been at the center of a research work with M. Binois (now with INRIA Sophia-Antipolis) and O. Roustant (now with INSA Toulouse) on the choice of the low-dimensional domain in high-dimensional global optimization via random embeddings, which as resulted in the article [1]. On a different note, a collaboration with Noémie
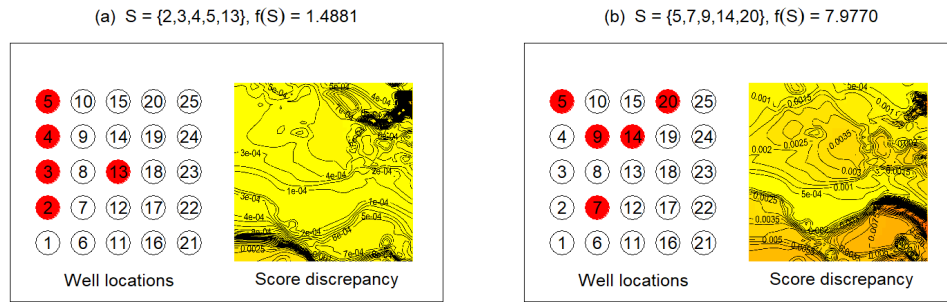
Figure 10: Score discrepancy map: location of selected wells (input S), score discrepancy landscape, and the spatial sum of score discrepancy objective function value $f(S)$ from [2].

Jaquier and Sylvain Calinon from Idiap's Robot Learning and Interaction (RLI) group has resulted in [5], where it is shown how to design in a multi-output GP framework a probabilistic model that encapsulates variability information inherited from on a preliminary Gaussian Mixture model learnt from demonstrations while being generative and naturally accommodating via-points and related constraints.

Computer experiments for the quantification and the reduction of uncertainties

Besides global optimization, UQOD has indeed also been focusing on sequential 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. More generally, the combination of GP modelling and stochastic simulation techniques have been investigated for quantifying and reducing uncertainties on sets, with example application in safety engineering with goals such as identifying sets of dangerous and/or safe configurations of a complex system and if possible providing some measures of confidence along with the estimate(s). Previously developed "asymmetric nested Monte Carlo" algorithms and their use for efficiently estimating orthant probabilities of high-dimensional Gaussian vectors have been instrumental to derive conservative set estimates on a neutronic criticality safety test case (from IRSN, the French Institut de Radioprotection et de Sûreté Nucléaire) and also to derive sequential design strategies dedicated to this class of conservative set estimation problem. The underlying long-standing collaboration with colleagues from Neuchâtel, CentraleSupélec and IRSN, has resulted in the article [3]. Current perspectives include the adaptation of such (conservative) set estimation strategies to larger scale inverse problem from geophysics (collaboration with UNIL), e.g., when uncertainties on the mass density field inside a volcano are to be reduced via well-chosen gravimetric measurements.

Key publications

[1] M. Binois, D. Ginsbourger, and O. Roustant, "On the choice of the low-dimensional domain for global optimization via random embeddings," in *J. Glob. Optim.* 76, 69–90 (2020).

[2] D. Ginsbourger and T. Krityakierne, "Kernels over sets of finite sets using RKHS embeddings, with application to Bayesian (combinatorial) optimization," *Accepted to AISTATS 2020*.

[3] D. Azzimonti, D. Ginsbourger, C. Chevalier, J. Bect, and Y. Richet, "Adaptive Design of Experiments for Conservative Estimation of Excursion Sets," in *Technometrics* (Published online: 23 Dec 2019).

[4] D. Armentano, J.-M. Azaïs, D. Ginsbourger, and J.R. León, "Conditions for the finiteness of the moments of the volume of level sets," in *Electron. Commun. Probab.*, 24:17 (2019).

[5] N. Jaquier, D. Ginsbourger, and S. Calinon, "Learning from demonstration with model-based Gaussian process," *Conference on Robot Learning (CoRL)*, 2019.

2.6 Genomics and Health Informatics

Overview

Head: Dr. Raphaëlle Luisier (Master of Science in Bioengineering and Biotechnology, EPFL, 2009 & PhD in Bioinformatics from the Basel University, 2013)

Group overview: The mass generation of high-content genomic, imaging and digital data has brought rapid advances in our understanding of cell biology and human diseases, and offers promising avenues for the field of precision medicine. A major obstacle revolves around delivering meaningful information from these data, requiring the deployment of a variety of methods such as statistics, machine learning and data visualisation in tight collaboration with clinicians. The Genomics and Health Informatics Group was created in 2019 to develop statistical and machine-learning methods to integrate genomic, clinical, and imaging data. The group aims to accelerate the diagnosis and improve the understanding and treatment of complex diseases such as neurodegenerative disorders in collaboration with clinical neuroscientists. The group has expertise in genomics, bioinformatics, RNA biology, neuroscience, data science, and data visualisation.

Key scientific outputs: The current lines of research include:

[1] The integration of longitudinal cellular imaging with genomic data to study how molecular biology shapes cellular morphology in neurodegenerative diseases;

[2] The development of machine learning methods to integrate longitudinal digital data with genomics to study how individual differences in motor and cognitive behaviours emerge and how such differences may confer vulnerability to neurodegenerative diseases.

In collaborating with clinicians and expert laboratories in molecular neuroscience and bioengineering, the group has access to unique high quality data including longitudinal RNA-sequencing (Patani laboratory, Francis Crick Institute, London) and time-lapse cellular imaging (Serio lab, Kings College, London) from motor neurone cultures obtained from healthy and sick donors.

Additional information and a list of projects are available from www.idiap.ch/genomics.

2.7   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 human-centered robotic applications in which the robots need to acquire skills from only few demonstrations and interactions, with strong generalization demands. It requires the development of intuitive 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 2019 was composed of 2 postdoctoral fellows, 7 PhD students and 2 visiting PhD/MSc students.

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 (e.g., assistive robots in I-DRESS), parts of us (e.g., prosthetic hands in TACT-HAND), or far away from us (e.g., 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.

Controllers relying on the minimal intervention principle
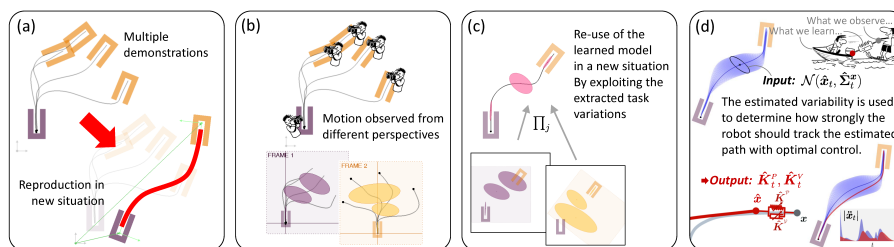
Figure 11: Task-parameterized model to synthesize movements in new situations. (a) Observation in different situations. (b) Probabilistic encoding in multiple coordinate systems. (c) Adaptation of the motion to new situations. (d) Model predictive controller exploiting variability and correlation information.

Model predictive control (MPC) is ubiquitous in robotics. Its core formulation can be extended to a wide range of problems, which has often been overlooked in robot learning. In particular, the most simple form of MPC already has advantage for motion synthesis and planning problems, where it can be combined elegantly with probabilistic representations of movements. This approach allows the retrieval of smooth and natural trajectories, by taking into account variation and coordination constraints. Instead of learning trajectories, the approach allows us to learn the underlying controllers to move 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 modify autonomously the tracking gains and compliance required to reproduce a task.

Geometry-aware learning and control

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
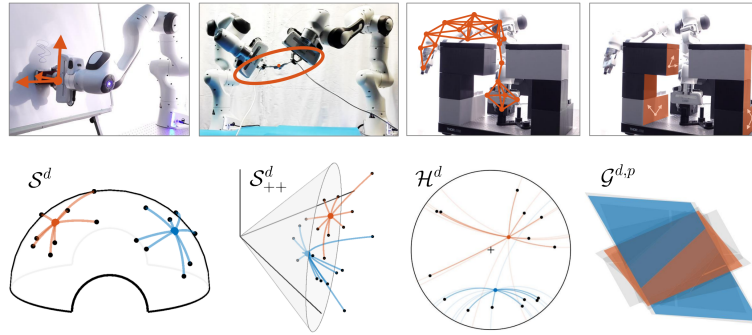
Figure 12: *Structured manifolds in robotics. $\mathcal{S}^3$ can be used to represent the orientation (unit quaternions). $\mathcal{S}^6_{++}$ can be used to represent manipulability ellipsoids (symmetric positive definite matrices). $\mathcal{H}^d$ can be used to represent trees, graphs and roadmaps. $\mathcal{G}^{d,p}$ can be used to represent subspaces (planes, nullspaces, projection operators).*

spatial covariances, 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 taking into account prior knowledge about these manifolds. It allows us to treat data of various forms in a unified manner, and to extend common optimization problems in robotics initially formulated in standard Euclidean spaces.

Tensor-variate regression in robotics applications
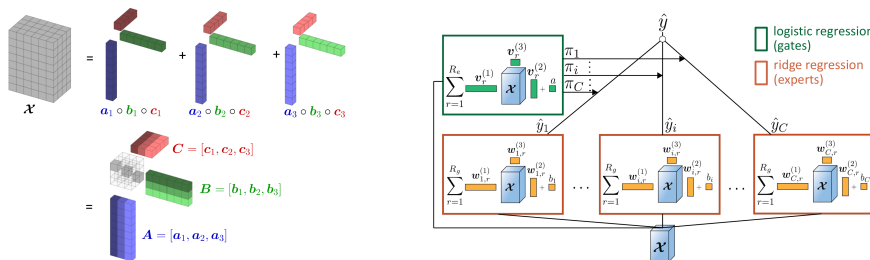
Figure 13: *Left: Tensor factorization. Right: Tensor-variate mixture of experts, with tensor regression as experts and tensor logistic regression as gating functions.*

In many robotics applications, data are sparse, either taking the form of demonstrations provided by the user, or experiences collected by the robot. The sensory data are typically organized as multidimensional arrays (arrays of sensors, multiple channels, time evolution of data, multiple coordinate systems, etc.). This led our group to investigate the use of tensor methods and related multilinear algebra techniques. Tensors are generalization of matrices to arrays of higher dimensions. Classical regression methods first transform these data into vectors, therefore ignoring the underlying structure, which typically leads to overfitting when only few training data are available. We address this issue through the use of expert models (product of experts, mixture of experts) relying on tensorial representations.

Key publications

[1] S. Calinon, "Mixture Models for the Analysis, Edition, and Synthesis of Continuous Time Series," Bouguila, N. and Fan, W. (eds.), *Mixture Models and Applications*, pp. 39-57. Springer, 2019.

[2] S. Calinon, and D. Lee, "Learning Control," Vadakkepat, P. and Goswami, A. (eds.), *Humanoid Robotics: a Reference*, pp. 1261-1312. Springer, 2019.

[3] E. Pignat, and S. Calinon, "Bayesian Gaussian Mixture Model for Robotic Policy Imitation," in *IEEE Robotics and Automation Letters* (RA-L), 4:4, 4452-4458, 2019.

2.8　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, 2008–2012,2012–2018)

Group overview: The Natural Language Understanding group works at the intersection of machine learning and natural language processing, with an emphasis on representation learning for the meaning of language, attention-based deep learning models, and structured prediction. We model summarization, abstraction (textual entailment), machine translation, knowledge extraction, syntactic and semantic structure, and lexical semantics, among other NLP problems. We develop deep learning models of the discovery and prediction of entities and their relations at multiple levels of representation for multiple tasks.

During 2019, the NLU group had the following members: the head of the group, two postdoctoral researchers, and 5 PhD students. Both postdoctoral researchers moved to positions elsewhere.

Key scientific outputs: During 2019, the work of the NLU group has included proposing a novel version of the Transformer deep learning architecture which maps graphs to graphs, applied to syntactic parsing, and representation learning for large numbers of output classes, applied to text categorization. Improved portability and transfer for models of natural language inference were demonstrated, both by using sentence embeddings with a direct entailment-based interpretation, and by training with novel bias-reduction methods. Also, the head-of-group helped submit two successful grants, an SNSF-FWO project on knowledge-base population, and an SNSF NCCR grant on the evolution of language, proposing biologically inspired deep learning models of syntactic processing.

Additional information and a list of projects are available from www.idiap.ch/nlp.

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 a in deep learning architecture which explicitly models the similarity between output classes. This can be done with output embedding 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 output-embedding 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-output-embedding classification has performed better. This effect is particularly marked with very large label sets, where our model also trains faster, in time independent of the label set size. We have also generalised this architecture to the case of decoding in machine translation, where there are a large number of possible next output words.

Neural Network Architectures for Graphs

Deep learning models based on self-attention, in particular Transformer, are revolutionising the state-of-the-art in many NLP tasks, particularly when combined with pretraining (e.g. BERT). We have developed a version of Transformer which maps graphs to graphs, instead of sequences to sequences. Given the nodes of a graph, such as the words of a sentence, Graph2Graph Transformer can input arbitrary graphs to the self-attention mechanism and can predict arbitrary graps with an attention-like output function. When combined with BERT pretraining, these models are amongst the best models for syntactic parsing, a core benchmark for structured prediction. We are currently investigating other structure prediction tasks, other NLP tasks which can benefit from structured inputs, and architectures which also predicting the nodes of the graph.

Key publications

[1] Nikolaos Pappas and James Henderson, "GILE: A Generalized Input-Label Embedding for Text Classification," accepted in *Transactions of the Association for Computational Linguistics (TACL)*, 2019.

[2] Alireza Mohammadshahi and James Henderson, "Graph-to-Graph Transformer for Transition-Based Dependency Parsing," ArXiv e-prints, arXiv:1911.03561 [cs.CL], 2019.

2.9   Computational Bioimaging

Overview

Head: Prof. Michael Liebling (MS, EPFL, 2000; PhD, EPFL 2004; postdoc, Caltech, 2004–2007; Assistant Prof (2007–2013), Associate Prof (2013–2017), Adjunct Prof (2017–) UC 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 deconvolution methods, and (ii) quantitative analysis of complex biological systems: motion-based image analysis, cell tracking, microscopic fluid flow estimation, and integration of multi-modality images.

In 2019, the Computational Bioimaging Group was composed of the head of the group, four PhD students, one postdoc, and an intern.

Key scientific outputs: Recent milestones include development of temporal super-resolution methods for sensitive fluorescence cameras, and approaches for quantitative imaging of heart development in animal models.

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

DeepFocus: a few-shot microscope slide auto-focus using a sample invariant CNN-based sharpness function

As part of the SNSF project COMP-Bio "Computational Biomicroscopy: Advanced Image Processing Methods to Quantify Live Biological Systems," we developed a microscopy autofocus (AF) method. AF methods are extensively used in biomicroscopy, for example to acquire timelapses, where the imaged objects tend to drift out of focus. AF algorithms determine an optimal distance by which to move the sample back into the focal plane. Current hardware-based methods require modifying the microscope and image-based algorithms either rely on many images to converge to the sharpest position or need training data and models specific to each instrument and imaging configuration. We proposed DeepFocus, an AF method we implemented as a plugin for a common microscopy driver (Micro-Manager) We characterized its Convolutional Neural Network (CNN)-based sharpness function, which we observed to be depth co-variant and sample-invariant. Sample invariance allows our AF algorithm to converge to an optimal axial position within as few as three iterations using a model trained once for use with a wide range of optical microscopes and a single instrument-dependent calibration stack acquisition of a flat (but arbitrary) textured object. We fperformance and low image numberoresee that this method will help limit photodamage during acquisitions with light-sensitive samples.

Temporal Super-Resolution Microscopy Using a Hue-Encoded Shutter

In a complementary research line of the COMP-bio project, we developed a computational imaging method for flexible imaging of color and high-speed events. We collaborated with the Institut de Recherche en Opthalmologie in Sion to demonstrate the applicability in biological samples and with Idiap's Robot learning group to characterize the method. Limited time-resolution in microscopy is an obstacle to many biological studies. Despite recent advances in hardware, digital cameras have limited operation modes that constrain frame-rate, integration time, and color sensing patterns. We proposed an approach to extend the temporal resolution of a conventional digital color camera by leveraging a multi-color illumination source. Our method allows imaging single-hue objects at increased frame-rate by trading spectral for temporal information (while retaining the ability to measure base hue). It also allows rapid switching to standard RGB acquisition. We evaluated the feasibility and performance of our method via experiments with mobile resolution targets.
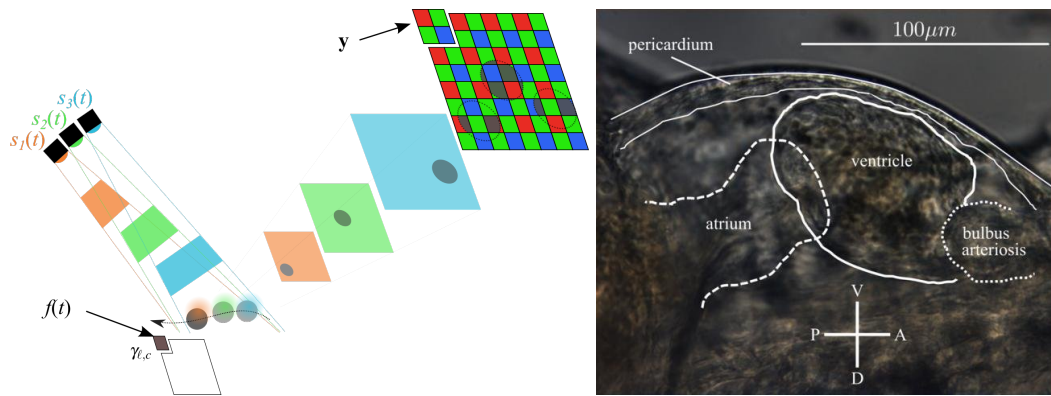
Figure 14: Hue-encoded shutter microscopy. (left) Acquisition setup with a moving sample illuminated by light sources that encode temporal information in different spectra and signals captured by a color sensor. Hue information is converted to higher frame-rate.(right) Heart of 4 dpf old zebrafish larva, imaged with temporal super-resolution in region of interest. (fig. adapted from C.J. et al. Biomed. Opt. Express, 2019

We observed a time-resolution increase by a factor 2.8 with a three-fold increase in temporal sampling rate. We further illustrate the use of our method to image the beating heart of a zebrafish larva, allowing the display of color or fast grayscale images. The method is particularly well-suited to extend the capabilities of imaging systems where the flexibility of rapidly switching between high frame rate and color imaging are necessary.

Reconstruction of image sequences from ungated and scanning-aberrated laser scanning microscopy images of the beating heart

Within the joint SNSF and French ANR project "liveheart: The cellular basis of cardiac development revealed by live imaging," our group developed methods for high-framerate microscopy of the beating heart. Fluorescence laser-scanning microscopy is a well-established imaging technique in biology, available in many imaging facilities to investigate structures within live animal embryos such as zebrafish. Laser scanning microscopes (LSM) are limited when used to study dynamic heart morphology or function. Despite their ability to resolve static cardiac structures, the fast motion of the beating heart introduces severe artifacts in the scanned images and gating the acquisitions to the heartbeat is difficult to implement on traditional microscopes. Our approach takes a set of images containing scanning aberrations, each triggered at an arbitrary time in the cardiac cycle, and assembles an image sequence that covers a single cardiac heartbeat. We characterize the performance of our method on synthetic data and further illustrated its applicability to experimental images acquired in live zebrafish larvae, and show that the reconstruction quality approaches that of fast, state-of-the-art microscopes. Our technique opens the possibility of using LSMs to carry out studies of cardiac dynamics, without the need for prospective gating or fast microscopes.

Key publications

[1] A. Shajkofci and M. Liebling, "DeepFocus: a few-shot microscope slide auto-focus using a sample invariant CNN-based Sharpness Function," in *IEEE International Symposium on Biomedical Imaging*, 2020, *in press*

[2] C. Jaques, E. Pignat, S. Calinon, and M. Liebling, "Temporal super-resolution microscopy using a hue-encoded shutter," in *Biomed. Opt. Express*, vol. 10, no. 09, pp. 4727–4741, Sept. 2019.

[3] O. Mariani, A. Ernst, N. Mercader, and M. Liebling, "Reconstruction of image sequences from ungated and scanning-aberrated laser scanning microscopy images of the beating heart," in *IEEE Trans. Comput. Imag.*, vol. 6, pp. 385–395, Oct. 2019.

2.10   Biometrics Security and Privacy

Overview

Head: Dr. Sébastien Marcel (PhD, University of Rennes, France, 2000; Visiting Professor, University of Cagliari, 2010; Lecturer, EPFL and UNIL, 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 (PAD): 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/), an *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 2019 was composed of 1 head of group, 3 research associates, 5 postdocs, 2 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, 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. 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 presentation attacks. The BSP group also investigated approaches for heterogeneous face recognition, vein recognition and more recently Deepfakes to assess its threat to face recognition.

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

Deepfakes: a New Threat to Face Recognition ?

It is becoming increasingly easy to automatically replace a face of one person in a video with the face of another person by using a pre-trained generative adversarial network (GAN). Recent public scandals call for automated ways to detect these Deepfake videos. To help developing such methods, we produced and analysed Deepfake videos. We showed that the state of the art face recognition systems based on Neural Networks are vulnerable to Deepfake videos, with about 90% false acceptance rates , which means methods for detecting Deepfake videos are necessary. By considering several baseline approaches, we found that audio-visual approach based on lip-sync inconsistency detection was not able to distinguish Deepfake videos. The best performing method, which is based on visual quality metrics and is often used in presentation attack detection domain, resulted in 8.97% equal error rate on high quality Deepfakes. Our experiments demonstrate that GAN-generated Deepfake videos are challenging for both face recognition systems and existing detection methods, and the further development of face swapping technology will make it even more so.

(a) Original Donor   (b) Original Target   (c) Face Swapped

Figure 15: Illustration of Deepfakes

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. 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.
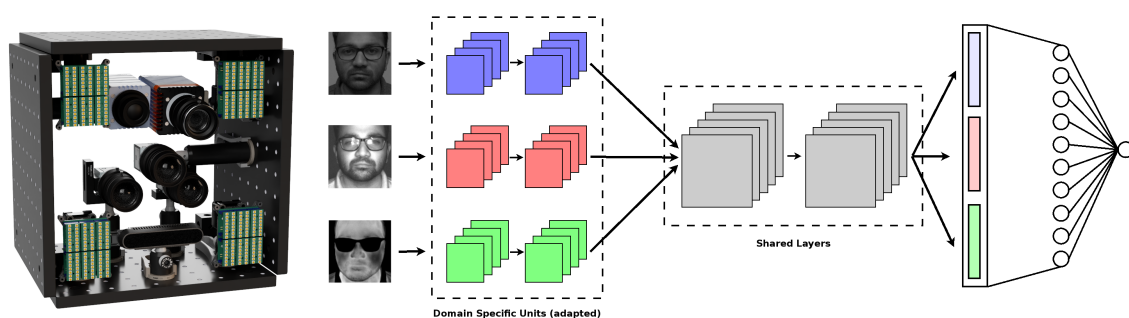
Figure 16: Illustration of multi-spectral face PAD device (left) and architecture (right).

We developed a multi-spectral face device (Figure 16) to capture synchronised, high speed and high resolution image sequences under different image domains (VIS, NIR, SWIR, Thermal and 3D). The main hypothesis is that bona fide samples are easier to discriminate from presentation attacks with the appropriate combination of image domains. We proposed a novel Deep Convolutional Neural Network architecture (Figure 16) to learn multi-spectral complementary information.

Heterogeneous face recognition

The task of Heterogeneous Face Recognition (Figure 17) 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 demonstrated that high level features of Deep Convolutional Neural Networks trained on visual spectra images are domain independent and can be used to encode faces sensed in different image domains.

Figure 17: 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).

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 18).
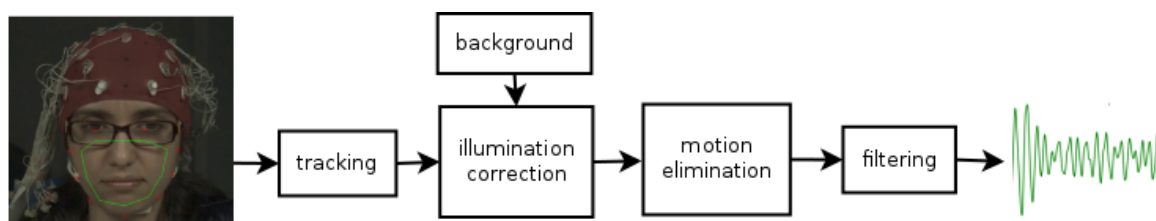
Figure 18: Illustration of remote photoplethysmography: colors from the video signal are filtered to produce an estimation of the heart beat signal.

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.

Swiss Centre for Biometrics Research and Testing

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 2018, the centre has developed over 3 directions:

- Improving the security of the BEAT platform: following an external IT security audit our engineers strengthened the platform. The platform is now used in multiple research projects such as ALLIES, LEARN-REAL and the major H2020 AI4EU project.

- Simplifying the BEAT platform: with the aim to be more attractive, our engineers developed a new feature to easily edit not only biometric experiments but also more general pattern recognition and AI-based experiments.

- Training: we were teaching the UniDistance Certificate in Biometrics and Privacy that is further developed for the AI Master program as one specific course on Biometrics. (https://distanceuniversity.ch/cas-biometrics/)

- Testing: we engaged with the FIDO Alliance (https://fidoalliance.org), focused on providing open and free authentication standards to help reduce the world's reliance on passwords, and became an accredited FIDO lab able to perform certification of biometrics products.

Key publications

[1] P. Korshunov and S. Marcel, "Vulnerability assessment and detection of Deepfake videos", *IAPR International Conference on Biometrics*, pp 258–264, 2019.

[2] G. Heusch and S. Marcel, "Remote Blood Pulse Analysis for Face Presentation Attack Detection", *Handbook of Biometric Anti-Spoofing (2nd Edition)*, Springer, Chapter 10, pp 267–289, 2019.

[3] K. Kotwal, S. Bhattacharjee and S. Marcel, "Multispectral Deep Embeddings As a Countermeasure To Custom Silicone Mask Presentation Attacks", *IEEE Transactions on Biometrics, Behavior, and Identity Science*, Vol 1(4), pp 238–255, 2019.

[4] A. George, Z. Mostaani, D. Geissenbuhler, O. Nikisins, A. Anjos and S. Marcel, "Biometric Face Presentation Attack Detection with Multi-Channel Convolutional Neural Network", *IEEE Transactions on Information Forensics and Security*, Vol 15, pp 42–55, 2019.

[5] T. de Freitas Pereira, A. Anjos and S. Marcel, "Heterogeneous Face Recognition Using Domain Specific Units", *IEEE Transactions on Information Forensics and Security*, Vol 14(7), pp 1803–1816, 2019.

2.11   Biosignal Processing

Overview

Head: Dr. André Anjos (BSc, MSc & PhD, Federal University of Rio de Janeiro, Brazil, 1999, 2001, 2006; Visiting Professor, State University of Bauru, 2015; Lecturer, EPFL, 2013-)

Group overview: Biosignals are signals from living beings and their analysis to support clinical practice and research. This group currently focuses on image (e.g. retinography, X-ray, computerized tomography), vital signs, clinical and multimodal data analysis for healthcare and related applications. Current trends in the field show refreshed interest on the use of machine learning techniques, complementing basic signal and sequence processing, all of which are key domains of research at Idiap. It leverages on Idiap's expertise on human subject handling, data acquisition, open science and data processing.

In 2019, the Biosignal Processing Group was composed of the head of the group, 1 master student and 1 intern.

Key scientific outputs: The group currently develops 4 thematic areas of research: computer-aided diagnosis of retinographies (eye fundus scans) via conventional 45° degree sensors, the prediction of adverse events in Latent Tuberculosis Infection (LTBI) treatment from clinical data, decompensation prediction from vital signs using portable sensors, and, finally, reproducibility in AI research. The milestones for 2020 include continued work in above cited areas, the development of new projects on chest X-ray for TB screening and computed tomography (CT) scans for the detection of rare vessel diseases, and finally continued participation in Idiap's AI master program. Additional funds have been request for the Swiss National Science Foundation for hiring three doctoral students.

Additional information and a list of projects are available from www.idiap.ch/bio-sig.

Semantic Segmentation for Medical Imaging

We identify and address two research gaps in the field of vessel segmentation for 2D eye-fundus imaging. The first focuses on the task of inference on high-resolution fundus images for which only a limited set of ground-truth data is publicly available. We showed that simple transformation techniques like rescaling, padding and cropping of combined lower-resolution source datasets to the resolution and spatial composition of a higher-resolution target dataset can be a surprisingly effective way to improve segmentation quality in unseen conditions. Our results show competitive performance on a set of common public retinal vessel datasets using a small and light-weight neural network. For HRF, the only very high-resolution dataset currently available, we reach new state-of-the-art (SOTA) performance by solely relying on training images from lower-resolution datasets. In comparison to previous works, we note that our approach achieves comparable performance without using any training images of the target-set, solely relying on the remaining publicly available datasets for training. Given the stark differences in color, illumination, brightness and spatial composition between datasets this is an encouraging result and illustrates the robustness of our approach.

The second gap concerns reported metrics in available reading material. We emphasized the need for a more rigorous and detailed focus on evaluation, and proposed a set of plots and metrics that give additional insights into model performance. We propose to use the standard deviation as a proxy for the confidence on the estimation of the average F1-score. We demonstrated via tables and plots how to take advantage of that information, throwing a new light over some published benchmarks.

Figure 19: Vessel segmentation for Retinography. Illustration of predicted vessel maps vs. ground truths for M2U-Net evaluated on the high-resolution HRF test-set. The source dataset was composed of low-resolution images. True positives, false positives and false negatives are displayed in green, blue and red respectively.

Reproducible Research

All our work is driven by fully reproducible framework[24]. More recently, we have been actively looking at the reproducibility of published work and how to lower the entrance barrier of publication readers. We argue it is insufficient, in most cases, to only publish software leading to results if original data remains inaccessible. Reproducibility should imply in the following characteristics: repeatability, share-ability, extensibility and stability, which is not guaranteed by most published material to date. We propose a software suite called Bob[25] that possesses such characteristics, demonstrating its flexibility to various tasks including Medical Image Segmentation, Biometric Person Recognition, Presentation Attack Detection, and Remote Photoplethysmography.

From another perspective, there are legitimate cases in which raw data leading to research conclusions cannot be published. Furthermore, in a growing number of use-cases, the availability of both software does not translate to an accessible reproducibility scenario. To bridge this gap, we built an open platform for research[26] in computational sciences related to pattern recognition and machine learning, to help on the development, reproducibility and certification of results obtained in the field.

Bob and BEAT are still active and support past and future work at Idiap and beyond. We conduct lectures to both master and graduate students about reproducibility in data science.

Key publications

[1] T. Laibacher and A. Anjos, "On the Evaluation and Real-World Usage Scenarios of Deep Vessel Segmentation for Retinography," Submitted to the 23rd International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), 2020.

[2] T. de Freitas Pereira, A. Anjos and S. Marcel, "Heterogeneous Face Recognition Using Domain Specific Units," in *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 7, pp. 1803-1816, July 2019.

---

[24]https://gitlab.idiap.ch/bob/bob.ip.binseg
[25]https://www.idiap.ch/software/bob
[26]https://www.idiap.ch/software/beat

2.12   Energy Informatics

Overview

Head: Dr. Jérôme Kämpf (BSc, University of Kent at Canterbury; MSc, University of Lausanne, 2001 and 2003; PhD, Ecole Polytechnique Fédérale de Lausanne, 2009)

Group overview: The Energy Informatics concepts are to exploit state-of-the-art Information and Communication Technologies to tackle global warming and climate change challenges. The aim is to increase the integration of renewable and distributed energy sources by making energy systems smarter, and to increase energy efficiency beyond what improvements at component level can achieve. In that vein, the Energy Informatics Group at Idiap researches into ways of simulating energy transition pathways with intelligent control and adjustment mechanisms of evolving shelters with retrofitting and use, renewable energy production and energy storage in a changing climate. Two key application areas are more deeply studied: the building automation with its adaptation to human behavior and the energy management with its indirect modeling of human activity to anticipate energy needs.

In 2019, the Energy Informatics Group was composed of 1 head of group, 1 postdoc, 2 exchange PhD students, 1 exchange MSc student and 1 intern.

Key scientific outputs: Two exchange PhD students from the National University of Singapore and the University of Aveiro (Portugal) spent each 6 months at Idiap. As a result of this collaboration, two conference papers relating to CO_2 monitoring for smart control in buildings, and one conference and one journal papers relating to surrogate models for the prediction of photovoltaic production of high-rise buildings, were published. In the meantime, an exchange MSc student from the University of Pisa spent 3 months at Idiap, resulting in two conference papers on reducing the performance gap using machine learning models for urban scale simulations. A post-doc from the University of Sydney (Australia) joined the group with the aim to instrument Idiap to monitor the sky radiance distribution and detect a trade-off between visual comfort, daylight penetration, electric lighting, heating and cooling energy needs. An intern worked on the evaluation of the urban heat island effect in the city of Fribourg through the calibration of the subtending physical models. The head of the group's PhD student at EPFL successfully defended publicly his thesis on "Automated Daylighting Control System based on Sky Luminance Monitoring and Lighting Computing" in May 2019, and filled 2 patents. In total, 5 journal papers and 9 conference papers have been published by the group in 2019.

Additional information and a list of projects are available from http://www.idiap.ch/energy.

Building automation

Building occupants tend to negatively perceive building automation as it may alter their comfort for the sake of energy savings. The main challenge identified is therefore to control (or suggest actions on) the building infrastructure in order to minimise the energy intake while maintaining a sufficient comfort for the occupants. Noteworthy, human comfort has multiple facets and depends on each individual's history and preferences. The use of smart sensors is a unique opportunity to learn and adapt the automation to the users.

Through the project INNO-EE, we continuously monitor CO_2 concentration in buildings as a proxy for the air quality and the air renewal rate. The results are analysed to evaluate the accuracy and optimal placement of the sensors, together with the necessity of giving feed-back to users regarding window opening actions for a healthier environment. On a longer term, the project is the premises for prediction algorithms of the air quality and energy consumption aiming at an optimal control for both low-tech and high-tech buildings.

Through the Mexican scholarship of a visiting post-doc, we study the trade-off between energy and comfort

in buildings. The energy consumption is considered for heating, cooling and lighting while the comfort considered as thermal and visual. The Idiap Research Institute's building is being set-up as a case study and demonstrator for lighting and blinds control within the framework of the IEA SHC Task 61 (task61.iea-shc.org) on Integrated Solutions for Daylighting and Electric Lighting: From component to user centered system efficiency.

Energy management

While urban areas cover about 2% of the planet surface, they consume 3/4 of the total resources. According to the Pareto rule, major energy-related efforts should target cities and their inhabitants. The efforts comprises the integration of renewable and distributed energy sources, which in dense urban centers is a challenge. In particular the mutual shading between buildings affect the active and passive solar performance.

Through the Singaporian scholarship of a visiting PhD student, we studied the urban morphology to determine the effectiveness of PV panels in the urban environment and hence the design of new urban developments. Statistical predictive models (from multi-linear regressions to more advanced machine learning algorithms) were built on the basis of detailed physical simulations of the radiation in the urban texture, allowing for an optimisation of the urban built form.

Human activity in buildings may have a strong impact on the overall energy consumption. An evaluation of the performance gap between the foreseen building energy consumption and the monitored one was approached by an exchange Master student from the University of Pisa. The features behind the differences were evaluated from a macroscopic point of view introducing new urban morphological indicators.

Through the projet ICU, we evaluated using physically-based simulations the Urban Heat Island effect in the city of Fribourg. The calibration of the building energy consumption for space heating is of prime importance to predict correct surface temperatures, likewise the soil stratigraphy and the vegetation presence in the urban environment. A database was set-up with the information available on the built and natural environments in Fribourg, and linked to the urban energy simulator CitySim to predict the hourly dynamics of the surface temperatures.

CREM - Centre de Recherches Energétiques et Municipales

We maintain a close scientific collaboration with the CREM (www.crem.ch), an actor specialized in the field of energy sustainability in urban areas. CREM has a large network of communes, public utilities and companies making it an essential partner for case-studies and dissemination. We shared an intern on the evaluation of the performance and automation of the Idiap Research Institute's building.

Key publications

[1] Y. Wu, T. Wang, E.S. Lee, J. Kämpf and J. -L. Scartezzini, "Split-pane electrochromic window control based on an embedded photometric device with real-time daylighting computing," in *Building and Environment*, vol. 161, 2019

[2] N. Mohajeri, A. Gudmundsson, G. Knuckler, D. Assouline, J. Kämpf and J. -L. Scartezzini, "A solar-based sustainable urban design: The effects of city-scale street-canyon geometry on solar access in Geneva," in *Applied Energy*, vol. 240, pp. 173-190, 2019

[3] Y. Wu, J. Kämpf and J.-L. Scartezzini, "Performance assessment of the BTDF data compression based on wavelet transforms in daylighting simulation," in *Solar Energy*, vol. 190, pp. 329-336, 2019

3 Researchers (in addition to permanent staff)

3.1 Research Associates

Lastname	Firstname	Nationality	Supervisor	Start	Estimated End
BHATTACHARJEE	Sushil	CH	Sébastien MARCEL	2019-01-01	2019-12-31
GEISSBUHLER	David	CH	Sébastien MARCEL	2018-06-01	2020-12-31
HEUSCH	Guillaume	CH	Sébastien MARCEL	2015-06-08	
IMSENG	David	CH	Hervé BOURLARD	2016-04-01	2024-11-30
KORSHUNOV	Pavel	EE	Sébastien MARCEL	2017-01-01	2020-06-30
LABHART	Florian	CH	Daniel GATICA-PEREZ	2017-11-01	2020-12-31
MADIKERI	Srikanth	IN	Petr MOTLICEK	2018-03-01	2020-07-31

3.2 Post-doctoral Scholars

Lastname	Firstname	Nationality	Supervisor	Start	Estimated End
ANTONELLO	Niccolò	IT	Philip GARNER	2019-04-15	2020-04-15
BABY	Deepak	IN	Hervé BOURLARD	2019-05-01	2020-05-01
CAO	Yuanzhouhan	CN	Jean-Marc ODOBEZ	2017-11-15	2019-02-15
DE FREITAS PEREIRA	Tiago	BR	Sébastien MARCEL	2019-03-01	2021-02-28
DEY	Subhadeep	IN	Petr MOTLICEK	2018-11-12	2020-04-30
EWERTON	Marco	DE	Jean-Marc ODOBEZ	2019-08-15	2020-08-14
GEORGE	Anjith	IN	Sébastien MARCEL	2017-08-21	2020-12-31
HALSTEAD	Michael	AU	Sébastien MARCEL	2018-09-01	2019-08-31
KHONGLAH	Banriskhem	IN	Hervé BOURLARD	2018-03-01	2020-01-07
KHOSRAVANI	Abbas	IR	Philip GARNER	2019-07-15	2020-07-15
KODRASI	Ina	AL	Hervé BOURLARD	2017-12-01	2019-12-31
KOTWAL	Ketan	IN	Sébastien MARCEL	2018-01-01	2020-07-31
KRIVOKUCA	Vedrana	HR	Sébastien MARCEL	2017-01-15	2021-01-15
LE	Nam	VN	Jean-Marc ODOBEZ	2019-04-01	2019-05-31
MASSÉ	Benoit	FR	Daniel GATICA-PEREZ	2019-03-01	2020-08-31
MURALIDHAR	Skanda	IN	Daniel GATICA-PEREZ	2018-12-01	2020-05-31
NGUYEN	Laurent	CH	Daniel GATICA-PEREZ	2015-04-01	2019-07-31
NIKISINS	Olegs	LV	Sébastien MARCEL	2014-07-01	2019-03-01
PAOLILLO	Antonio	IT	Sylvain CALINON	2018-01-15	2019-08-21
PAPPAS	Nikolaos	GR	James HENDERSON	2016-02-01	2019-08-31
PARIDA	Shantipriya	IN	Petr MOTLICEK	2019-02-01	2022-01-31
PRASAD	Ravi	IN	Mathew MAGIMAI DOSS	2018-10-01	2021-10-31
RAM	Dhananjay	IN	Michael LIEBLING	2019-07-01	2020-02-28
REKABSAZ	Navid	IR	James HENDERSON	2018-03-01	2019-09-30
SAMUI	Suman	IN	Petr MOTLICEK	2019-09-01	2019-11-30
SARFJOO	Saeed	IR	Sébastien MARCEL	2018-01-01	2020-06-30
SILVERIO	João	PT	Sylvain CALINON	2019-07-01	2020-06-30
VILLAMIZAR	Michael	ES	Jean-Marc ODOBEZ	2016-11-01	2020-10-31

3.3 PhD Students

Lastname	Firstname	Nationality	Supervisor	Start	Estimated End
AMINIAN	Bozorgmehr	CH	Jean-Marc ODOBEZ	2018-04-01	2019-09-15
CHAVDAROVA	Tatjana	MK	François FLEURET	2014-12-01	2020-07-31
COURDIER	Evann	FR	François FLEURET	2019-03-01	2023-02-28
DE FREITAS PEREIRA	Tiago	BR	Sébastien MARCEL	2014-07-01	2019-02-28
DESPRÈS	Nicolas	FR	Jean-Marc ODOBEZ	2018-02-01	2019-04-30
DUBAGUNTA	Pavankumar	IN	Mathew MAGIMAI DOSS	2017-04-01	2021-03-31
FRITSCH	Julian	DE	Mathew MAGIMAI DOSS	2018-07-01	2022-06-30
GAUTIER	Athenais	FR	David GINSBOURGER	2018-11-01	2022-10-31
GIRGIN	Hakan	TR	Sylvain CALINON	2018-09-01	2022-08-31
HAJIBABAEI	Mahdi	IR	Petr MOTLICEK	2019-06-11	2019-11-30
HE	Weipeng	CN	Jean-Marc ODOBEZ	2016-06-01	2020-05-31
HERMANN	Enno	DE	Mathew MAGIMAI DOSS	2018-07-01	2022-06-30
JANBAKHSHI	Parvaneh	IR	Hervé BOURLARD	2018-02-01	2022-01-31
JAQUES	Christian	CH	Michael LIEBLING	2016-04-01	2020-03-31
JAQUIER	Noémie	CH	Sylvain CALINON	2016-08-02	2020-08-01
KABIL	Selen	TR	Hervé BOURLARD	2018-05-01	2022-04-30
KARIMI MAHABADI	Rabeeh	IR	James HENDERSON	2018-10-01	2022-09-30
KATHAROPOULOS	Angelos	GR	François FLEURET	2017-03-01	2021-02-28
KULAK	Thibaut	FR	Sylvain CALINON	2017-12-01	2021-11-30
LE	Nam	VN	Jean-Marc ODOBEZ	2015-01-15	2019-05-31
LEMBONO	Teguh	ID	Sylvain CALINON	2018-07-23	2022-07-23
MAI	Florian	DE	James HENDERSON	2018-10-01	2022-09-30
MARELLI	François	BE	Michael LIEBLING	2018-10-01	2022-09-30
MARFURT	Andreas	CH	James HENDERSON	2018-11-01	2022-10-31
MARIANI	Olivia	CH	Michael LIEBLING	2016-06-01	2020-05-31
MARTINEZ GONZALEZ	Angel	MX	Jean-Marc ODOBEZ	2016-11-15	2020-11-15
MEEGAHAPOLA	Iakmal	LK	Daniel GATICA-PEREZ	2019-06-24	2023-06-23
MICULICICH	Lesly	PE	James HENDERSON	2016-03-01	2020-09-30
MOHAMMADI	Amir	IR	Sébastien MARCEL	2016-01-15	2020-01-31
MOHAMMADSHAHI	Alireza	IR	James HENDERSON	2019-09-01	2022-08-31
MUCKENHIRN	Hannah	FR	Mathew MAGIMAI DOSS	2015-07-13	2019-08-31
PHAN	Trung	VN	Daniel GATICA-PEREZ	2018-09-01	2020-09-30
PIGNAT	Emmanuel	CH	Sylvain CALINON	2016-03-01	2020-08-31
RAM	Dhananjay	IN	Michael LIEBLING	2015-01-01	2019-06-30
SAJADMANESH	Sina	IR	Daniel GATICA-PEREZ	2019-05-15	2023-05-15
SCHNELL	Bastian	DE	Philip GARNER	2017-05-01	2021-04-30
SHAJKOFCI	Adrian	CH	Michael LIEBLING	2016-09-20	2020-09-20
SHETTY	Suhan	IN	Sylvain CALINON	2019-06-03	2023-06-03
SIEGFRIED	Rémy	CH	Jean-Marc ODOBEZ	2017-01-01	2020-12-31
SIVAPRASAD	Prabhu	IN	François FLEURET	2018-11-01	2022-10-31
SRINIVAS	Suraj	IN	François FLEURET	2017-03-01	2021-02-28
TONG	Sibo	CN	Philip GARNER	2016-05-01	2020-04-30
TORNAY	Sandrine	CH	Mathew MAGIMAI DOSS	2016-08-01	2020-07-31

TRAVELLETTI	Cédric	CH	David GINSBOURGER	2018-11-01	2022-10-31
TROUSSARD	Martin	FR	Sylvain CALINON	2019-01-02	2023-01-02
VYAS	Apoorv	IN	Hervé BOURLARD	2018-07-15	2022-07-15
YU	Yu	CN	Jean-Marc ODOBEZ	2015-07-15	2019-12-31
ZHAN	Qingran	CN	Petr MOTLICEK	2018-10-01	2020-03-31

3.4 Interns

Lastname	Firstname	Nationality	Supervisor	Start	Estimated End
BARRAS	Léonard	CH	Michael LIEBLING	2019-07-22	2020-01-31
BEHJATI	Melika	IR	James HENDERSON	2019-12-01	2020-08-31
BOGHETTI	Roberto	IT	Jérôme KÄMPF	2019-01-08	2019-03-29
BUATHONG	Poompol	TH	David GINSBOURGER	2019-03-01	2019-08-31
CHAUDHARY	Sapana	IN	Sylvain CALINON	2019-05-01	2019-05-01
COURDIER	Evann	FR	François FLEURET	2018-10-22	2019-02-28
DI BLASI	Anthony	CH	Jérôme KÄMPF	2019-07-01	2020-12-31
GAO	Xiao	CN	Sylvain CALINON	2019-10-01	2020-09-30
GOMEZ ALANIS	Alejandro	ES	Mathew MAGIMAI DOSS	2019-09-01	2019-11-30
HAJIBABAEI	Mahdi	IR	Petr MOTLICEK	2019-02-10	2019-06-10
JEANNINGROS	Loïc	FR	Philip GARNER	2018-11-01	2019-04-30
LAIBACHER	Tim	DE	André RABELLO DOS ANJOS	2019-03-20	2019-09-20
LOKIETKO	Jaroslaw	PL	Sylvain CALINON	2019-07-01	2019-12-31
MATOBA	Kyle	US	François FLEURET	2019-10-14	2020-04-30
MOSTAANI	Zohreh	IR	Sébastien MARCEL	2018-04-09	2019-04-09
NGUYEN	Huy	VN	Sébastien MARCEL	2019-10-10	2020-02-07
ORRÙ	Giulia	IT	Sébastien MARCEL	2019-11-15	2020-02-14
PAGLIARDINI	Matteo	FR	François FLEURET	2019-12-01	2020-02-29
PATEL	Kumar	IN	James HENDERSON	2018-10-12	2019-02-28
PEREIRA DE OLIVEIRA	Rui	PT	Jérôme KÄMPF	2019-03-06	2019-09-06
POON	Kin Ho	GB	Jérôme KÄMPF	2019-03-28	2019-10-13
SAINT-SUPERY	Santiago	FR	Daniel GATICA-PEREZ	2019-09-15	2020-02-29
SHARMA	Shivam	IN	Mathew MAGIMAI DOSS	2018-10-01	2019-03-31
VITEK	Radovan	CZ	Hervé BOURLARD	2019-06-01	2019-08-31
WANG	Lei	CN	Hervé BOURLARD	2019-09-01	2020-08-31
YELA BELLO	Jenny Paola	CO	Daniel GATICA-PEREZ	2019-11-18	2020-06-30

3.5 Visitors

Lastname	Firstname	Nationality	Supervisor	Start	Estimated End
BASURTO DAVILA	Chantal	AU	Jérôme KÄMPF	2019-10-01	2020-03-31
FRIEDLI	Lea	CH	David GINSBOURGER	2019-01-22	2020-12-31
KÖPRÜLÜ	Cevahir	TR	Michael LIEBLING	2019-02-18	2019-05-31
MAALOUF	Eliane	LB	David GINSBOURGER	2018-11-01	2021-11-30
MENDOZA	Viviana	CU	Mathew MAGIMAI DOSS	2019-06-24	2019-07-07
MOHAMMADSHAHI	Alireza	IR	James HENDERSON	2019-02-25	2019-08-31
POZO PEREZ	Ada Laura	ES	Daniel GATICA-PEREZ	2019-02-15	2019-11-15
VAES	Clément	BE	Sylvain CALINON	2019-09-01	2020-05-30
VILLATORO TELLO	Esaú	MX	Hervé BOURLARD	2019-09-01	2020-08-31

4   Active and Granted Projects in 2019

An overview of the projects that have been active during the year 2019 is presented in Section 4.1. The projects are grouped in three categories, namely National Research Projects, European and International Research Projects and Industry-oriented Projects.

Section 4.2 presents the list of projects accepted during 2019 but starting in the following year.

4.1   Projects in Progress during 2019

4.1.1   National Research Projects

[1]   Name   **ADIVA** (Automatic analysis of verbal and non-verbal behavior and provision of feedback in video selection interviews)
        Funding   SNF - Division I
        Coordinator   University of Neuchâtel
        Duration   2018.12.01 - 2020.05.31
        Partner(s)   Idiap Research Institute, University of Lausanne

[2]   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

[3]   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
        Partner(s)   -

[4]   Name   **COMPBIO** (Computational biomicroscopy: advanced image processing methods to quantify live biological systems)
        Funding   SNF - Division II
        Coordinator   Idiap Research Institute
        Duration   2018.04.01 - 2022.03.31
        Partner(s)   -

[5]   Name   **DEVEL-IA** (Formation « Développeurs spécialisés en Intelligence Artificielle » selon le modèle de formation continue duale postgrade)
        Funding   SECO
        Coordinator   Idiap Research Institute
        Duration   2018.05.01 - 2019.06.30
        Partner(s)   Formation universitaire à Distance

[6] Name   **DOMAT** (On-demand Knowledge for Document-level Machine Translation)
     Funding   SNF - Division II
     Coordinator   Idiap Research Institute
     Duration   2018.10.01 - 2022.09.30
     Partner(s)   HES-SO Vaud

[7] Name   **DUSK2DAWN** (Characterizing Youth Nightlife Spaces, Activities, and Drinks)
     Funding   SNF - Sinergia
     Coordinator   Idiap Research Institute
     Duration   2017.07.01 - 2020.12.31
     Partner(s)   University of Zurich, La Trobe University

[8] Name   **FLOSS** (Flexible Linguistically-guided Objective Speech aSessment)
     Funding   Hasler Foundation
     Coordinator   Idiap Research Institute
     Duration   2017.03.01 - 2020.03.31
     Partner(s)   -

[9] Name   **HEALTHVLOGGING** (Social media culture and the (re)shaping of health-related practices by Youtubers)
     Funding   SNF Spark
     Coordinator   University of Lausanne
     Duration   2019.12.01 - 2020.11.30
     Partner(s)   Idiap Research Institute

[10] Name   **HEAP** (Human-Guided Learning and Benchmarking of Robotic Heap Sorting)
     Funding   SNF - ERA NET
     Coordinator   University of Lincoln
     Duration   2019.04.01 - 2022.03.31
     Partner(s)   Idiap Research Institute, Istituto Italiano di Tecnologia, Institut de Recherche en Informatique et en Automatique, Technische Universitaet Wien

[11] Name   **ICU** (Ilots de chaleur en ville de Fribourg : identification, anticipation et stratégie d'adaptation et de valorisation)
     Funding   OFEN
     Coordinator   HES-So Fribourg
     Duration   2019.01.01 - 2020.12.31
     Partner(s)   Idiap Research Institute

[12] Name   **I-DRESS** (Assistive Interactive robotic system for support in DRESSing)
     Funding   SNF - ERA NET
     Coordinator   Idiap Research Institute
     Duration   2015.12.01 - 2019.02.28
     Partner(s)   -

[13] Name   **INTREPID** (Automated interpretation of political and economic policy documents: Machine learning using semantic and syntactic information)
      Funding   SNF - Sinergia
      Coordinator Graduate Institute of International and Development Studies
      Duration   2019.01.01 - 2022.12.31
      Partner(s)   Idiap Research Institute

[14] Name   **IPEQ** (Uncertainty quantification and efficient design of experiments for data- and simulation-driven inverse problem solving)
      Funding   SNF - Division II
      Coordinator Idiap Research Institute
      Duration   2018.11.01 - 2022.08.31
      Partner(s)   University of California at Davis

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

[16] Name   **LAB_FIDO** (Création d'un laboratoire de test biométrique agréé)
      Funding   Loterie Romande
      Coordinator Idiap Research Institute
      Duration   2018.12.01 - 2019.11.30
      Partner(s)   -

[17] Name   **LAOS** (Learning Representations of Abstraction in Text)
      Funding   SNF - Division II
      Coordinator Idiap Research Institute
      Duration   2018.11.01 - 2022.09.30
      Partner(s)   -

[18] Name   **LEARN-REAL** (LEARNing physical manipulation skills with simulators using REAListic variations)
      Funding   SNF - ERA NET
      Coordinator Idiap Research Institute
      Duration   2019.04.01 - 2022.03.31
      Partner(s)   Ecole Centrale de Lyon, Istituto Italiano di Tecnologia

[19] 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 - 2019.11.30
      Partner(s)   Idiap Research Institute, University of Bern, École Polytechnique Paris

[20] Name   **LOIS** (Leveraging on-device smartphone inference to address resistance to participate in social surveys)
    Funding   UNIL/EPFL
    Coordinator   University of Lausanne
    Duration   2019.01.01 - 2019.12.31
    Partner(s)   Idiap Research Institute, École Polytechnique Fédérale de Lausanne

[21] Name   **MASS** (Multilingual Affective Speech Synthesis)
    Funding   SNF - Division II
    Coordinator   Idiap Research Institute
    Duration   2017.05.01 - 2020.09.30
    Partner(s)   -

[22] Name   **MEMUDE** (Multi-view Detection with Metric-learning for Deep Network Fusion)
    Funding   Hasler Foundation
    Coordinator   Idiap Research Institute
    Duration   2017.06.01 - 2019.11.30
    Partner(s)   -

[23] Name   **MOSPEEDI** (Motor Speech Disorders: characterizing phonetic speech planning and motor speech programming/execution and their impairments)
    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

[24] Name   **MPM** (Multimodal People Monitoring)
    Funding   Idiap-CSEM Program
    Coordinator   Idiap Research Institute
    Duration   2018.02.01 - 2019.12.31
    Partner(s)   Centre Suisse d'Electronique et de Microtechnique

[25] Name   **ODESSA** (Online Diarization Enhanced by recent Speaker identification and Sequential learning Approaches)
    Funding   SNF - ANR
    Coordinator   Centre National de La Recherche Scientifique
    Duration   2016.03.01 - 2019.10.31
    Partner(s)   Idiap Research Institute, Eurecom

[26] Name   **PHASER-QUAD** (Parsimonious Hierarchical Automatic Speech Recognition and Query Detection)
    Funding   SNF - Division II
    Coordinator   Idiap Research Institute
    Duration   2016.10.01 - 2019.09.30
    Partner(s)   -

[27] Name   **REAPPS** (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

[28] Name   **ROBOCLETTE** (Roboclette (Robot racleur))
     Funding   Etat du Valais - SHE
     Coordinator   Idiap Research Institute
     Duration   2019.05.01 - 2019.12.31
     Partner(s)   -

[29] Name   **ROSALIS** (Robot skills acquisition through active learning and social inter-
                     action strategies)
     Funding   SNF - Division II
     Coordinator   Idiap Research Institute
     Duration   2018.04.01 - 2022.03.31
     Partner(s)   -

[30] Name   **SHISSM** (Sparse and hierarchical Structures for Speech Modeling)
     Funding   SNF - Division II
     Coordinator   Idiap Research Institute
     Duration   2018.01.01 - 2021.12.31
     Partner(s)   -

[31] Name   **SMILE** (Scalable Multimodal sign language Technology for sIgn language
                     Learning and assessmEnt)
     Funding   SNF - Sinergia
     Coordinator   Idiap Research Institute
     Duration   2016.03.01 - 2020.10.31
     Partner(s)   University of Surrey, University of Applied Sciences of Special Needs Edu-
                     cation

[32] 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 - 2020.05.31
     Partner(s)   Idiap Research Institute, Universitaet Bielefeld

[33] Name   **TIPS** (Towards Integrated processing of Physiological and Speech signals)
     Funding   SNF Division II
     Coordinator   Idiap Research Institute
     Duration   2019.11.01 - 2023.10.31
     Partner(s)   Centre Suisse d'Electronique et de Microtechnique, Coaching & Modera-
                     tion

[34] Name   **TOT** (Trust Over Time)
     Funding   EPFL
     Coordinator   École Polytechnique Fédérale de Lausanne
     Duration   2019.10.01 - 2020.09.30
     Partner(s)   Idiap Research Institute, Radio Television Suisse

[35] Name   **UNITS** (Unified Speech Processing Framework for Trustworthy Speaker Recognition)
     Funding   SNF - Division II
     Coordinator   Idiap Research Institute
     Duration   2015.07.01 - 2019.06.30
     Partner(s)   -

[36] Name   **VERIFAKE** (Vulnerability assessment and detection of Deepfake videos)
     Funding   Hasler Foundation
     Coordinator   Idiap Research Institute
     Duration   2019.06.01 - 2020.05.31
     Partner(s)   -

4.1.2   European and International Research Projects

[1] 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

[2]   Name   **AI4EU** (A European AI On Demand Platform and Ecosystem)
        Funding   H2020-RIA-ICT
        Coordinator   Thales Services SA
        Duration   2019.01.01 - 2021.12.31
        Partner(s)   Idiap Research Institute, ABB AS, AGI Research SRO, Allianz SE, Atos Spain SA, Aristotle University of Thessaloniki, Blumorpho SAS, Budapest University of Technology and Economics, Bureau de Recherches Géologiques et Minières, Barcelona Supercomputing Center, CARTIF Foundation, Commissariat à l'énergie atomique et aux énergies alternatives, CINECA - Consorzio Interuniversitario, Consiglio Nationale delle Ricerche, Centre National de la Recherche Scientifique, Agencia Estatal Consejo Superior De Investigaciones Cientifícas, National Center for Sicentific Research "Demokritos", German Research Center for Artificial Intelligence, German Aerospace Center, EIT Digital, Eötvös Loránd University, European Organisation for Security, FundingBox Research, Fondazione Bruno Kessler, Fraunhofer Gesellschaft, France Digitale, Foundation for Research and Technology – Hellas, Forum Virium Helsinki, Grassroots Arts and Research, France Hub IA, Industrial Data Space e. V., Institut Mines-Télécom-IMT, Institut National de Recherche en Informatique et Automatique, Associacao Do Instituto Superior Tecnico Para A Investigacao E Desenvolvimento, Centre for Research and Technology Hellas, Jožef Stefan Institute, Karlsruhe Institute of Technology, Know-Center Gmbh Research Center for Datadriven Business & Big data Analytics, University of Leeds, Loupe 16, Università degli studi di Roma "La Sapienza", Norwegian University of Science and Technology, National University of Ireland Galway, Office National d'Etudes et Recherches Aérospatiales, Orange SA, Örebro University, PG WConsulting, Université Paris I Panthéon-Sorbonne, QWANT, Siemens AG, SAP SE, Smile, Smart Rural, Simula Research Laboratory, Thales Alenia Space, Thomson Licensing, Telenor ASA, Tilde SIA, Thales Research & Technology France, Technische Universität Berlin, Delft University of Technology, Centre for Intelligent Technologies, Technical University of Munich, Technical University of Vienna, Alma Mater Studiorum – University of Bologna, University College Cork, University of Coimbra, Université Grenoble Alpes, Unilever UK, National and Kapodistrian University of Athens, Universitat Politecnica de Catalunya, Technical University of Madrid, Universita Degli Studi Di Siena, Sorbonne University, Ca' Foscari University of Venice, Vrije Universiteit Brussel, Women in AI, Wavestone,

[3]   Name   **ATCO2** (Automatic collection and processing of voice data from air-traffic communications)
        Funding   H2020-CSJU
        Coordinator   Idiap Research Institute
        Duration   2019.11.01 - 2022.02.28
        Partner(s)   Brno University of Technology, OpenSky Network, ReplayWell, Romagna Tech, Evaluations and Language Resources Distribution Agency, University of Saarland

[4]   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

[5] Name   **COLLABORATE** (Co-production CeLL performing Human-Robot Collaborative AssEmbly)
     Funding   H2020-RIA-DT
     Coordinator Aristotle University Of Thessaloniki
     Duration   2018.10.01 - 2021.09.30
     Partner(s)   Idiap Research Institute, Arcelik A.S., Association pour la Recherche et le Développement des méthodes et processes, ASTI Mobile Robotics, Blue Ocean Robotics APS, Centre for Research and Technology Hellas, Centro Ricerche Fiat SCPA, Jozef Stefan Institute, Katholieke Universiteit Leuven, Kolektor, University of Patras, Pratt & Whitney Rzeszów, Universita Degli Studi di Genova

[6] 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

[7] Name   **MUMMER** (MultiModal Mall Entertainment Robot)
     Funding   H2020-RIA-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

[8] Name   **ROXANNE** (Real time network, text, and speaker analytics for combating organized crime)
     Funding   H2020-SU-SEC
     Coordinator Idiap Research Institute
     Duration   2019.09.01 - 2022.08.31
     Partner(s)   Trilateral Research LTD, Brno University of Technology, Phonexia s.r.o., SAIL LABS Technology GmbH, Capgemini Technology Services, The International Criminal Police Organization, Saarland University, KENTRO MELETON ASFALEIAS, Gottfried Wilhelm Leibniz Universitaet Hannover, Università Cattolica del Sacro Cuore – Transcrime, AEGIS IT RESEARCH UG, AIRBUS Defence and Space SAS (Innovation Coordinator), Police of Czech Republic, Romanian Minister of Interior, Lithuanian Forensic Science Centre, Police Service of Northern Ireland, ADITESS Advanced Integrated Technology Solutions & Services LTD, Ministry of Interior Croatia, Netherlands Forensic Institute, Internet of Things applications and Multi-Layer development, Ministry Of Public Security - Israel National Police, Hellenic Police, An Garda Síochána

[9] Name   **SARAL** (Summarization and domain-Adaptive Retrieval of Information Across Languages)
     Funding   USA IARPA
     Coordinator   University of Southern California
     Duration   2017.10.01 - 2020.07.31
     Partner(s)   Idiap Research Institute, Massachusetts Institute of Technology, Raytheon Company, Reenselaer Polytechnic Institute, University of Massachusetts Amherst, Northeastern University

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

[11] Name   **SUMMA** (Scalable Understanding of Multilingual Media)
     Funding   H2020-RIA-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

[12] 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

[13] 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 desenvolvimento Em Lisboa, Interuniversitair Micro-Electronica Centrum Imec Vzw, Stichting Het Nederlands Kanker Instituut - Antoni Van Leeuwenhoek Ziekenhuis, Universitaet Augsburg

[14] Name   **TESLA** (An Adaptive Trust-based e-assesment System for Learning)
     Funding   H2020-IA-ICT
     Coordinator Fundacio Per A La Universitat Oberta de Catalunya
     Duration   2016.01.01 - 2019.03.31
     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

[15] Name   **WENET** (The Internet of US)
     Funding   H2020-RIA-FETPROACT
     Coordinator University of Trento
     Duration   2019.01.01 - 2022.12.31
     Partner(s)   Idiap Research Institute, Aalborg University, Amrita Vishwa Vidyapeetham, Ben-Gurion University of the Negev, University of Tübingen, Instituto Potosino de Investigacion Cientifica y Tecnologica, Jilin University, London School of Economics and Political Science, Martel GmbH, National University of Mongolia, Open University of Cyprus, U-Hopper SRL, Universidad Catolica Nuestra Senora de La Ascuncion

4.1.3 Industry-oriented Projects

[1] Name   **AADES** (Adaptive and Asynchronous Detection and Segmentation)
    Funding   Armasuisse
    Coordinator   Idiap Research Institute
    Duration   2018.10.01 - 2019.09.30
    Partner(s)   -

[2] Name   **ADATIS** (Adatis - Optimisation du Contrôle Qualité)
    Funding   Fondation The Ark
    Coordinator   Adatis
    Duration   2018.10.01 - 2019.05.31
    Partner(s)   Idiap Research Institute

[3] Name   **ADVANCE** (Augmented dialogue tool based on verbal and non-verbal behavior computing)
    Funding   Innosuisse
    Coordinator   Idiap Research Institute
    Duration   2019.03.01 - 2020.08.31
    Partner(s)   CM Profiling, HES-SO Fribourg

[4] Name   **AMICLEAR** (Détection de fibres d'AMIante par CLassification d'images de microscopie Electronique pour l'Analyse de Résidus)
Funding   Fondation The Ark
Coordinator   Idiap Research Institute
Duration   2019.08.01 - 2020.04.01
Partner(s)   Amiscan

[5] Name   **BOAT** (Automated Braces generation for Orthopaedic Anatomical Treatment of fractures)
Funding   Innosuisse
Coordinator   Idiap Research Institute
Duration   2019.04.05 - 2020.08.04
Partner(s)   HES-SO Fribourg, Swibrace SA

[6] Name   **COBHOOK** (COB'HOOK)
Funding   Innosuisse
Coordinator   Idiap Research Institute
Duration   2019.05.01 - 2020.10.31
Partner(s)   Richemont

[7] Name   **DAHL** (Domain Adaptation via Hierarchical Lexicons)
Funding   Industrial
Coordinator   Idiap Research Institute
Duration   2019.11.01 - 2020.10.31
Partner(s)   Swisscom

[8] Name   **INNO-SIMPLYHOME** (Improving indoor environment, energy consupmtion and comfort in existing buildings)
Funding   Innosuisse
Coordinator   Idiap Research Institute
Duration   2019.10.01 - 2020.03.31
Partner(s)   Cleveron AG

[9] Name   **RISE** (Rich Interpersonal Skill analytics for rEcruitment)
Funding   Innosuisse
Coordinator   Idiap Research Institute
Duration   2018.05.01 - 2020.04.30
Partner(s)   University of Lausanne

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

[11]   Name   **SM2** (Extracting Semantic Meaning from Spoken Material)
        Funding   Innosuisse
        Coordinator   Idiap Research Institute
        Duration   2018.11.01 - 2020.04.30
        Partner(s)   Recapp

[12]   Name   **SPORTPROFILING** (Sport Profiling)
        Funding   Fondation The Ark
        Coordinator   Idiap Research Institute
        Duration   2019.08.01 - 2020.06.01
        Partner(s)   Action Types Swiss Sàrl, ProKey Coach

[13]   Name   **UNICITY** (3D scene understanding through machine learning to secure entrance zones)
        Funding   Innosuisse
        Coordinator   Idiap Research Institute
        Duration   2017.03.01 - 2019.10.31
        Partner(s)   Hes-So Fribourg, Fastcom Technology SA

[14]   Name   **USP** (Unique Stability Plates: Advanced Aluminium Solution for High Precision Milling)
        Funding   Innosuisse
        Coordinator   Idiap Research Institute
        Duration   2019.05.01 - 2020.11.01
        Partner(s)   Allega GmbH, Constellium Valais SA, Hes-so Valais

[15]   Name   **VIEW-2** (Visibility Improvement for Events Webcasting)
        Funding   Innosuisse
        Coordinator   Hes-So Valais
        Duration   2016.08.01 - 2019.09.30
        Partner(s)   Idiap Research Institute, Klewel

4.2   Projects Awarded in 2019 and Starting in the following year

[1]   Name   **ADEL** (Automatic Detection of Leadership from Voice and Body)
        Funding   UNIL
        Coordinator   University of Lausanne
        Duration   2020.06.01 - 2021.05.31
        Partner(s)   Idiap Research Institute, IMD Switzerland, Ecole Polytechnique Fédérale de Lausanne

[2]   Name   **AMS_SPONSORSHIP** (Sensor Fusion and Active Sensing for World-View Understanding)
        Funding   Industrial
        Coordinator   Idiap Research Institute
        Duration   2019.09.01 - 2024.08.31
        Partner(s)   ams

[3] Name   **CANDY** (ContActless finger veiN recognition and presentation attack De-
                   tection on-the-flY)
    Funding   Innosuisse
    Coordinator   Global ID SA
    Duration   2020.03.01 - 2022.02.28
    Partner(s)   Idiap Research Institute

[4] Name   **CODIMAN** (A future that works: Cobotics, digital skills and the re-
                   humanization of the workplace)
    Funding   SNF - NRP77
    Coordinator   Berner Fachhochschule
    Duration   2020.03.01 - 2024.02.29
    Partner(s)   Idiap Research Institute

[5] Name   **CORTI** (Computational Reduction for Training and Inference)
    Funding   SNF - Division II
    Coordinator   Idiap Research Institute
    Duration   2020.03.01 - 2022.02.28
    Partner(s)   -

[6] Name   **ENERMAPS** (Open Source Tools to Share, Compare, and Reuse Low-
                   Carbon Energy Data)
    Funding   H2020
    Coordinator   Centre de recherches energetiques et municipales
    Duration   2020.04.01 - 2022.03.31
    Partner(s)   Idiap Research Institute, Zentrum Für Energiewirtschaft Und Umwelt (E-
                   THINK), Accademia Europea di Bolzano, OpenAire Make, Revolve Media,
                   Technische Universitaet Wien

[7] Name   **EVOLANG** (The Origins and Future of Language)
    Funding   SNF - NCCR
    Coordinator   University of Zurich
    Duration   2020.04.01 - 2024.03.31
    Partner(s)   Idiap Research Institute, Ecole Polytechnique Fédérale de Lausanne, Ei-
                   dgenoessische Technische Hochschule Zuerich, University of Basel, Uni-
                   versity of Fribourg, University of Geneva, University of Lausanne, University
                   of Neuchâtel, Zurich University of the Arts

[8] Name   **GAZESENSESCREEN** (GazeSense Screen)
    Funding   Fondation The Ark
    Coordinator   Eyeware
    Duration   2020.01.01 - 2020.11.30
    Partner(s)   Idiap Research Institute

[9]   Name   **HAAWAII** (Highly Automated Air Traffic Controller Workstations with Artificial Intelligence Integration)
        Funding   H2020
        Coordinator   Deutsches Zentrum Fuer Luft und Raumfahrt Ev
        Duration   2020.06.01 - 2022.11.30
        Partner(s)   Idiap Research Institute, Austro Control, Croatia Control, Isavia OHF, NATS, Brno University of Technology

[10] Name   **ICARUS** (Innovative AppRoach to Urban Security)
        Funding   H2020
        Coordinator   Forum Européen Pour La Sécurité Urbaine
        Duration   2020.06.01 - 2024.05.31
        Partner(s)   Idiap Research Institute, Ethical and Legal Plus S.L., Erasmus Universiteit Rotterdam, Eurocircle Association, Fachhochschule Salzburg GmbH, Globalz SA, KENTRO MELETON ASFALEIAS, University of Leeds, Lisbon Municipal Police, Makesense, Commune de Nice, Panteion University Of Social And Political Sciences, Riga Municipal Police, City of Rotterdam, Landeshauptstadt Stuttgart, City of Torino, University of Salford

[11] Name   **MALAT** (Machine Learning for Air Traffic)
        Funding   Innosuisse
        Coordinator   Idiap Research Institute
        Duration   2020.03.01 - 2023.02.28
        Partner(s)   SkySoft ATM

[12] Name   **MHCSS** (Meaningful Human Control of Security Systems – Aligning International Humanitarian Law with Human Psychology)
        Funding   SNF - NRP77
        Coordinator   University of St.Gallen
        Duration   2020.05.01 - 2024.04.30
        Partner(s)   Idiap Research Institute, University of Zurich

[13] Name   **NAST** (Neural Architectures for Speech Technology)
        Funding   SNF - Division II
        Coordinator   Idiap Research Institute
        Duration   2020.02.01 - 2024.01.31
        Partner(s)   -

[14] Name   **NKBP** (Deep Learning Models for Continual Extraction of Knowledge from Text)
        Funding   SNF - Division II Lead Agency
        Coordinator   Idiap Research Institute
        Duration   2020.06.01 - 2024.05.31
        Partner(s)   Katholieke Universiteit Leuven

[15] Name   **SEWS2** (Smart Early Waning Score System for in and out-hospital care via anomaly detection)
      Funding   Innosuisse
      Coordinator   Idiap Research Institute
      Duration   2019.11.01 - 2020.05.31
      Partner(s)   Vtuls Sàrl

[16] Name   **TRESPASS-ETN** (TRaining in Secure and PrivAcy-preserving biometricS)
      Funding   H2020-MSCA
      Coordinator   Eurecom
      Duration   2020.01.01 - 2023.12.31
      Partner(s)   Idiap Research Institute, Hochschule Darmstadt, Chalmers Tekniska Hoegskola AB, Katholieke Universiteit Leuven, Rijksuniversiteit Groningen, Universidad Autónoma de Madrid

[17] Name   **WHEAT-ADVISOR** (Site-adapted variety choice for winter wheat in Switzerland)
      Funding   Federal Office for Agriculture
      Coordinator   Swiss Granum
      Duration   2020.01.01 - 2024.06.30
      Partner(s)   Idiap Research Institute, Agroscope

5   List of Publications in 2019

5.1   Books and Book Chapters

[1]   A. Anjos, P. Tome, and S. Marcel, "An introduction to vein presentation attacks and detection," in *Handbook of Biometric Anti-Spoofing*, S. Marcel, M. Nixon, J. Fierrez, and N. Evans, Eds., 2nd, Springer International Publishing, 2019, ch. 18.

[2]   D. Berio, F. F. Leymarie, and S. Calinon, "Interactive generation of calligraphic trajectories from gaussian mixtures," in *Mixture Models and Applications*, N. Bouguila and W. Fan, Eds., Springer, 2019, pp. 23–38.

[3]   S. Bhattacharjee, A. Mohammadi, A. Anjos, and S. Marcel, "Recent advances in face presentation attack detection," in *Handbook of Biometric Anti-Spoofing*, ser. Advances in Computer Vision and Pattern Recognition, S. Marcel, M. Nixon, J. Fierrez, and N. Evans, Eds., 2nd, Springer, Apr. 2019, ch. 10.

[4]   S. Calinon, "Learning from demonstration (programming by demonstration)," in *Encyclopedia of Robotics*, M. H. Ang, O. Khatib, and B. Siciliano, Eds., Springer, 2019.

[5]   S. Calinon, "Mixture models for the analysis, edition, and synthesis of continuous time series," in *Mixture Models and Applications*, N. Bouguila and W. Fan, Eds., Springer, 2019, pp. 39–57.

[6]   S. Calinon and D. Lee, "Learning control," in *Humanoid Robotics: A Reference*, P. Vadakkepat and A. Goswami, Eds., Springer, 2019, pp. 1261–1312.

[7]   I. Chingovska, A. Mohammadi, A. Anjos, and S. Marcel, "Evaluation methodologies for biometric presentation attack detection," in *Handbook of Biometric Anti-Spoofing*, S. Marcel, M. Nixon, J. Fierrez, and N. Evans, Eds., 2nd, Springer International Publishing, 2019, ch. 20.

[8]   G. Heusch and S. Marcel, "Remote blood pulse analysis for face presentation attack detection," in *Handbook of Biometric Anti-Spoofing*, ser. Advances in Computer Vision and Pattern Recognition, S. Marcel, M. Nixon, J. Fierrez, and N. Evans, Eds., 2nd, Springer, Apr. 2019, ch. 10.

[9]   I. Himawan, S. Madikeri, P. Motlicek, M. Cernak, S. Sridharan, and C. Fookes, "Voice presentation attack detection using convolutional neural networks," in *Handbook of Biometric Anti-Spoofing*, ser. Advances in Computer Vision and Pattern Recognition, 2nd, Springer, Apr. 2019, ch. 17, pp. 391–415.

[10] V. Krivokuca, M. Gomez-Barrero, S. Marcel, C. Rathgeb, and C. Busch, "Towards measuring the amount of discriminatory information in finger vein biometric characteristics using a relative entropy estimator," in *Handbook of Vascular Biometrics*, A. Uhl, C. Busch, S. Marcel, and R. Veldhuis, Eds., Springer Open, 2019, ch. 17, pp. 507–525.

[11] V. Krivokuca and S. Marcel, "On the recognition performance of biohash-protected finger vein templates," in *Handbook of Vascular Biometrics*, A. Uhl, C. Busch, S. Marcel, and R. Veldhuis, Eds., Springer Open, 2019, ch. 15, pp. 465–480.

5.2   Articles in Scientific Journals

[1]   N. Antonello, E. De Sena, M. Moonen, A. P. Naylor, and T. van Waterschoot, "Joint acoustic localiza-
      tion and dereverberation through plane wave decomposition and sparse regularization," *IEEE/ACM
      Transactions on Audio, Speech, and Language Processing*, vol. 27, no. 12, pp. 1893–1905, Dec.
      2019.

[2]   D. Armentano, J.-M. Azaïs, D. Ginsbourger, and J. R. León, "Conditions for the finiteness of the
      moments of the volume of level sets," *Electronic Communications in Probability*, vol. 24, no. 17,
      2019.

[3]   D. Azzimonti, D. Ginsbourger, C. Chevalier, J. Bect, and Y. Richet, "Adaptive design of experiments
      for conservative estimation of excursion sets," *Technometrics*, 2019.

[4]   D. Azzimonti, D. Ginsbourger, J. Rohmer, and D. Idier, "Profile extrema for visualizing and quantifying
      uncertainties on excursion regions. application to coastal flooding," *Technometrics*, vol. 61, no. 4,
      pp. 474–493, 2019.

[5]   J. Bect, F. Bachoc, and D. Ginsbourger, "A supermartingale approach to gaussian process based
      sequential design of experiments," *Bernoulli*, vol. 25, no. 4A, pp. 2883–2919, 2019.

[6]   M. Binois, D. Ginsbourger, and O. Roustant, "On the choice of the low-dimensional domain for global
      optimization via random embeddings," *Journal of Global Optimization*, 2019.

[7]   K. Chatzilygeroudis, A. Vassiliades, F. Stulp, S. Calinon, and J. B. Mouret, "A survey on policy search
      algorithms for learning robot controllers in a handful of trials," *IEEE Trans. on Robotics*, 2019.

[8]   M. Ewerton, O. Arenz, G. Maeda, D. Koert, Z. Kolev, M. Takahashi, and J. Peters, "Learning trajectory
      distributions for assisted teleoperation and path planning," *Frontiers in Robotics and AI*, vol. 6, p. 89,
      2019.

[9]   J. Ferris, C. Puljević, F. Labhart, A. Winstock, and E. Kuntsche, "The role of sex and age on pre-
      drinking: An exploratory international comparison of 27 countries," *Alcohol and Alcoholism*, vol. 54,
      no. 4, 378–385, Jul. 2019.

[10] T. de Freitas Pereira, A. Anjos, and S. Marcel, "Heterogeneous face recognition using domain spe-
      cific units," *IEEE Transactions on Information Forensics and Security*, p. 13, Feb. 2019.

[11] A. George, Z. Mostaani, D. Geissenbuhler, O. Nikisins, A. Anjos, and S. Marcel, "Biometric face
      presentation attack detection with multi-channel convolutional neural network," *IEEE Transactions
      on Information Forensics and Security*, 2019.

[12] I. Havoutis and S. Calinon, "Learning from demonstration for semi-autonomous teleoperation," *Au-
      tonomous Robots*, vol. 43, no. 3, pp. 713–726, 2019.

[13] G. Heusch, T. de Freitas Pereira, and S. Marcel, "A comprehensive experimental and reproducible
      study on selfie biometrics in multistream and heterogeneous settings," *IEEE Transactions on Bio-
      metrics, Behavior and Identity Science*, 2019.

[14] C. Jaques, E. Pignat, S. Calinon, and M. Liebling, "Temporal super-resolution microscopy using
      a hue-encoded shutter," *Optical Society of America Biomedical Optics Express*, vol. 10, no. 09,
      pp. 4727–4741, Sep. 2019.

[15] K. Kotwal, S. Bhattacharjee, and S. Marcel, "Multispectral deep embeddings as a countermeasure to
      custom silicone mask presentation attacks," *IEEE Transactions on Biometrics, Behavior, and Identity
      Science*, 2019.

[16] K. Kotwal, Z. Mostaani, and S. Marcel, "Detection of age-induced makeup attacks on face recognition
      systems using multi-layer deep features," *IEEE Transactions on Biometrics, Behavior, and Identity
      Science*, p. 11, 2019.

[17]   T. Laibacher and A. Anjos, *On the evaluation and real-world usage scenarios of deep vessel seg-mentation for funduscopy*, 2019. arXiv: `1909.03856 [cs.CV]`.

[18]   N. Le and J.-M. Odobez, "Improving speech embedding using crossmodal transfer learning with audio-visual data," *Multimedia Tools and Applications*, vol. 78, no. 11, pp. 15 681–15 704, Jan. 2019.

[19]   T. S. Lembono, F. Suarez-Ruiz, and Q.-C. Pham, "Scalar: Simultaneous calibration of 2-d laser and robot kinematic parameters using planarity and distance constraints," *IEEE Transactions on Automation Science and Engineering*, vol. 16, no. 4, R. Hytowitz, Ed., pp. 1971–1979, Oct. 2019.

[20]   O. Mariani, A. Ernst, N. Mercader, and M. Liebling, "Reconstruction of image sequences from un-gated and scanning-aberrated laser scanning microscopy images of the beating heart," *IEEE Trans-actions on Compuational Imaging*, 2019.

[21]   A. Martínez-González, M. Villamizar, O. Canévet, and J.-M. Odobez, "Efficient convolutional neural networks for depth-based multi-person pose estimation," *IEEE Transactions on Circuits and Systems for Video Technology*, 2019.

[22]   N. Mohajeri, A. Gudmundsson, G. Knuckler, D. Assouline, J. Kämpf, and J. L. Scartezzini, "A solar-based sustainable urban design: The effects of city-scale street-canyon geometry on solar access in geneva, switzerland," *Applied Energy*, vol. 240, pp. 173–190, Apr. 2019.

[23]   S. Okada, L. S. Nguyen, O. Aran, and D. Gatica-Perez, "Modeling dyadic and group impressions with inter-modal and inter-person features," *ACM Transactions on Multimedia Computing, Commu-nications, and Applications*, vol. 15, no. 1, Jan. 2019.

[24]   D. Palaz, M. Magimai.-Doss, and R. Collobert, "End-to-end acoustic modeling using convolutional neural networks for hmm-based automatic speech recognition," *Speech Communication*, vol. 108, pp. 15–32, Apr. 2019.

[25]   N. Pappas and J. Henderson, "Gile: A generalized input-label embedding for text classification," *Transactions of the Association for Computational Linguistics (TACL)*, 2019.

[26]   T.-T. Phan, F. Labhart, and D. Gatica-Perez, "My own private nightlife: Understanding youth personal spaces from crowdsourced video," *Proc. ACM Hum.-Comput. Interact*, vol. 3, no. 189, Nov. 2019.

[27]   T.-T. Phan, S. Muralidhar, and D. Gatica-Perez, "Drinks & crowds: Characterizing alcohol consump-tion through crowdsensing and social media," *Journal and Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT '19)*, Jun. 2019.

[28]   E. Pignat and S. Calinon, "Bayesian gaussian mixture model for robotic policy imitation," *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 4452 –4458, Oct. 2019.

[29]   G. Pirot, T. Krityakierne, D. Ginsbourger, and P. Renard, "Contaminant source localization via bayesian global optimization," *Hydrology and Earth System Sciences*, vol. 23, pp. 351–369, 2019.

[30]   S. Ruiz-Correa, I. Ruiz-Correa, C. Olmos Carrillo, F. A. Rendón-Huerta, B. Ramirez-Salazar, L. S. Nguyen, and D. Gatica-Perez, "Mi casa es su casa? examining airbnb hospitality exchange practices in a developing economy," *ACM Transactions on Social Computing*, vol. 2, no. 1, Feb. 2019.

[31]   N. Schütz, H. Saner, B. Rudin, A. Botros, B. Pais, V. Santschi, P. Buluschek, D. Gatica-Perez, P. Ur-wyler, L. Marchal-Crespo, R. M. Müri, and T. Nef, "Validity of pervasive computing based continuous physical activity assessment in community-dwelling old and oldest-old," *Scientific Reports*, vol. 9, no. 9662, Jul. 2019.

[32]   J. Silverio, S. Calinon, L. Rozo, and D. G. Caldwell, "Learning task priorities from demonstrations," *IEEE Transactions on Robotics*, vol. 35, no. 1, pp. 78–94, 2019.

[33]   O. Stanesby, F. Labhart, P. Dietze, C. Wright, and E. Kuntsche, "The contexts of heavy drinking: A systematic review of the combinations of context-related factors associated with heavy drinking occasions," *PLOS ONE*, vol. 14, no. 7, p. 29, Jul. 2019.

[34] A. K. Tanwani and S. Calinon, "Small variance asymptotics for non-parametric online robot learning," *International Journal of Robotics Research (IJRR)*, vol. 38, no. 1, pp. 3–22, 2019.

[35] S. Tornay and M. Magimai.-Doss, "Subunits inference and lexicon development based on pairwise comparison of utterances and signs," *Information*, vol. 10, p. 298, 2019.

[36] Y. Wu, J. Kämpf, and J. L. Scartezzini, "Automated eye-sight venetian blinds based on an embedded photometric device with real-time daylighting computing," *Applied Energy*, vol. 252, Oct. 2019.

[37] Y. Wu, J. Kämpf, and J. L. Scartezzini, "Daylighting simulation for external venetian blinds based on hdr sky luminance monitoring with matrix algebraic approach," *Energy Procedia*, vol. 158, pp. 2677–2682, Feb. 2019.

[38] Y. Wu, J. Kämpf, and J. L. Scartezzini, "Performance assessment of the btdf data compression based on wavelet transforms in daylighting simulation," *Solar Energy*, 2019.

[39] Y. Wu, T. Wang, E. S. Lee, J. Kämpf, and J. L. Scartezzini, "Split-pane electrochromic window control based on an embedded photometric device with real-time daylighting computing," *Building and Environment*, 2019.

5.3 PhD Theses

[1]   P. Dighe, "Sparse and low-rank modeling for automatic speech recognition," PhD thesis, EPFL, 2019.

[2]   T. de Freitas Pereira, "Learning how to recognize faces in heterogeneous environments," PhD thesis, Ecole Polytechnique Federale de Lausanne, Feb. 2019.

[3]   N. Le, "Multimodal person recognition in audio-visual streams," PhD thesis, EPFL, 2019.

[4]   H. Muckenhirn, "Trustworthy speaker recognition with minimal prior knowledge using neural networks," PhD thesis, Ecole polytechnique fédérale de Lausanne (EPFL), Switzerland, 2019.

[5]   S. Muralidhar, "Social sensing methods for analysis of dyadic hospitality encounters," PhD thesis, EPFL, Jan. 2019.

[6]   D. Ram, "Language independent query by example spoken term detection," PhD thesis, École Polytechnique Fédérale de Lausanne, Nov. 2019.

[7]   Y. Wu, J. Kämpf, and J. L. Scartezzini, "Automated daylighting control system based on sky luminance monitoring and lighting computing," PhD thesis, EPFL, 2019.

5.4 Articles in Conference Proceedings

[1] R. Boghetti, F. Fantozzi, J. Kämpf, G. Mutani, G. Salvadori, and V. Todeschi, "Building energy models with morphological urban-scale parameters: A case study in turin," in *Proceedings of 4th Building Simulation Applications Conference - BSA 2019*, Jun. 2019.

[2] E. Boschee, J. Barry, J. Billa, M. Freedman, T. Gowda, C. Lignos, C. Palen-Michel, M. Pust, B. Khonglah, S. Madikeri, J. May, and S. Miller, "Saral: A low-resource cross-lingual domain-focused information retrieval system for effective rapid document triage," in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: System Demonstrations. 2019*, 2019, pp. 19–24.

[3] T. Chavdarova, G. Gidel, F. Fleuret, and S. Lacoste-Julien, "Reducing noise in gan training with variance reduced extragradient," in *Proceedings of the international conference on Neural Information Processing Systems*, 2019.

[4] N. Dawalatabad, S. Madikeri, H. A. Murthy, and C. C. Sekhar, "Incremental transfer learning in two-pass information bottleneck based speaker diarization system for meetings," in *Proceedings of ICASSP 2019*, May 2019, pp. 6291–6295.

[5] S. Dey, P. Motlicek, T. Bui, and F. Dernoncourt, "Exploiting semi-supervised training through a dropout regularization in end-to-end speech recognition," in *Proc. of Interspeech 2019*, 2019.

[6] S. P. Dubagunta, S. H. Kabil, and M. Magimai.-Doss, "Improving children speech recognition through feature learning from raw speech signal," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2019.

[7] S. P. Dubagunta and M. Magimai.-Doss, "Segment-level training of anns based on acoustic confidence measures for hybrid hmm/ann speech recognition," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2019.

[8] S. P. Dubagunta and M. Magimai.-Doss, "Using speech production knowledge for raw waveform modelling based styrian dialect identification," in *Proceedings of Interspeech*, 2019.

[9] S. P. Dubagunta, B. Vlasenko, and M. Magimai.-Doss, "Learning voice source related information for depression detection," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2019.

[10] M. Ewerton, G. Maeda, D. Koert, Z. Kolev, M. Takahashi, and J. Peters, "Reinforcement learning of trajectory distributions: Applications in assisted teleoperation and motion planning," in *IEEE International Conference on Intelligent Robots and Systems*, 2019.

[11] J. Fritsch, S. Wankerl, and E. Nöth, "Automatic diagnosis of alzheimer's disease using neural network language models," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2019.

[12] D. Gatica-Perez, J.-I. Biel, D. Labbe, and N. Martin, "Discovering eating routines in context with a smartphone app," in *Ubicomp/Iswc'19 Adjunct: Proceedings Of The 2019 Acm International Joint Conference On Pervasive And Ubiquitous Computing And Proceedings Of The 2019 Acm International Symposium On Wearable Computers*, London, Sep. 2019, pp. 422–429.

[13] D. Gatica-Perez, D. Santani, J.-I. Biel, and T.-T. Phan, "Social multimedia, diversity, and global south cities: A double blind side," in *Proc. ACM Workshop on Fairness, Accountability, and Transparency in Multimedia (FAT/MM)*, Nice, Oct. 2019.

[14] A. George and S. Marcel, "Deep pixel-wise binary supervision for face presentation attack detection," in *International Conference on Biometrics*, 2019.

[15] M. Grisard, P. Motlicek, W. Allouchi, M. Baeriswyl, A. Lazaridis, and Q. Zhan, "Spoken language identification using language bottleneck features," in *Proceedings of TSD*, 2019.

[16] D. von Gunten, J. Rager, J. Kämpf, F. Kuchler, and F. Poumadère, "Retrofitting, district heating and energy storage: Neighborhood energy planning," in *Journal of Physics: Conference Series*, vol. 1343, IOP Publishing Ltd, Nov. 2019.

[17] W. He, P. Motlicek, and J.-M. Odobez, "Adaptation of multiple sound source localization neural networks with weak supervision and domain-adversarial training," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Brighton, United Kingdom, May 2019, pp. 770–774.

[18] P. Janbakhshi, I. Kodrasi, and H. Bourlard, "Pathological speech intelligibility assessment based on the short-time objective intelligibility measure," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2019.

[19] P. Janbakhshi, I. Kodrasi, and H. Bourlard, "Spectral subspace analysis for automatic assessment of pathological speech intelligibility," in *Proceedings of Interspeech*, Sep. 2019.

[20] C. Jaques, L. Bapst-Wicht, D. F. Schorderet, and M. Liebling, "Multi-spectral widefield microscopy of the beating heart through post-acquisition synchronization and unmixing," in *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, Venice, Italy, Apr. 2019, pp. 1382–1385.

[21] C. Jaques and M. Liebling, "Generalized temporal sampling with active illumination in optical microscopy," in *Proceeding of the SPIE Conference Optics and Photonics, Wavelets and Sparsity XVIII*, SPIE, vol. 11138, San Diego, California, United States: SPIE, Aug. 2019.

[22] N. Jaquier, D. Ginsbourger, and S. Calinon, "Learning from demonstration with model-based gaussian process," in *Conference on Robot Learning*, Oct. 2019.

[23] N. Jaquier, L. Rozo, S. Calinon, and M. Buerger, "Bayesian optimization meets riemannian manifolds in robot learning," in *Conference on Robot Learning*, Oct. 2019.

[24] A. Katharopoulos and F. Fleuret, "Processing megapixel images with deep attention-sampling models," in *Proceedings of International Conference on Machine Learning*, 2019.

[25] M. Kleinert, H. Helmke, G. Siol, h. Ehr, D. Klakow, M. Singh, P. Motlicek, K. Christian, C. Aneta, and H. Petr, "Adaptation of assistant based speech recognition to new domains and its acceptance by air traffic controllers," in *Proceedings of the 2nd International Conference on Intelligent Human Systems Integration (IHSI 2019): Integrating People and Intelligent Systems*, San Diego, California, USA, Feb. 2019, pp. 820 –826.

[26] I. Kodrasi and H. Bourlard, "Super-gaussianity of speech spectral coefficients as a potential biomarker for dysarthric speech detection," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, May 2019.

[27] P. Korshunov, M. Halstead, D. Castan, M. Graciarena, M. McLaren, B. Burns, A. Lawson, and S. Marcel, "Tampered speaker inconsistency detection with phonetically aware audio-visual features," in *International Conference on Machine Learning*, ser. Synthetic Realities: Deep Learning for Detecting AudioVisual Fakes, Best paper award in ICML workshop "Synthetic Realities: Deep Learning for Detecting AudioVisual Fakes", Jul. 2019.

[28] P. Korshunov and S. Marcel, "Vulnerability assessment and detection of deepfake videos," in *IAPR International Conference on Biometrics*, Jun. 2019.

[29] P. Korshunov and S. Marcel, "Vulnerability of face recognition to deep morphing," in *International Conference on Biometrics for Borders*, Oct. 2019.

[30] S. Madikeri, S. Dey, and P. Motlicek, "A bayesian approach to inter-task fusion for speaker recognition," in *In Proceedings of ICASSP 2019*, Brighton, ENGLAND, May 2019, pp. 5786–5790.

[31] rabeeh karimi mahabadi, J. lin, and V. Cevher, "A learning-based framework for quantized compressed sensing," in *A Learning-Based Framework for Quantized Compressed Sensing*, 2019.

[32] F. Mai, L. Galke, and A. Scherp, "Cbow is not all you need: Combining cbow with the compositional matrix space model," in *International Conference on Learning Representations*, New Orleans, Louisiana, USA, 2019.

[33] F. Marelli, B. Schnell, H. Bourlard, T. Dutoit, and P. N. Garner, "An end-to-end network to synthesize intonation using a generalized command response model," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brighton, United Kingdom: IEEE, May 2019, pp. 7040–7044.

[34] O. Mariani, K. G. Chan, A. Ernst, N. Mercader, and M. Liebling, "Virtual high-framerate microscopy of the beating heart via sorting of still images," in *2019 IEEE 16th International Symposium on Biomedical Imaging*, Apr. 2019, pp. 312–315.

[35] L. Miculicich, M. Marone, and H. Hassan, "Selecting, planning, and rewriting: A modular approach for data-to-document generation and translation," in *WNGT EMNLP*, 2019.

[36] A. Mohammadshahi, K. Aberer, and R. Lebret, "Aligning multilingual word embeddings for cross-modal retrieval task," in *Proceedings of the Second Workshop on Fact Extraction and VERification (FEVER)*, Hong Kong, China: Association for Computational Linguistics, Nov. 2019, pp. 27–33.

[37] P. Motlicek, H. Hermansky, S. Madikeri, A. Prasad, and S. Ganapathy, "Am-fm decomposition of speech signal: Applications for speech privacy and diagnosis," in *11th International workshop on Models and Analysis of Vocal Emissions for Biomedical Applications*, Universita Degli Studi Firenze, Firenze, Italy, Dec. 2019.

[38] H. Muckenhirn, V. Abrol, M. Magimai.-Doss, and S. Marcel, "Understanding and visualizing raw waveform-based cnns," in *Proceedings of Interspeech*, 2019.

[39] E. Naboni, M. Meloni, C. Makey, and J. Kämpf, "The simulation of mean radiant temperature in outdoor conditions: A review of architectural tools calculation assumptions," in *Proceedings of Building Simulation 2019: 16th Conference of IBPSA*, Sep. 2019.

[40] A. Nanchen and P. N. Garner, "Empirical evaluation and combination of punctuation prediction models applied to broadcast news," in *Proceedings of 2019 IEEE International Conference on Acoustics, Speech, and Signal Processing*, May 2019.

[41] J. Nembrini, J. Kämpf, M. Pappinutto, and D. Lalanne, "A smart luminaire in an office environment: Impact on light distribution, user interactions and comfort," in *Journal of Physics: Conference Series*, vol. 1343, IOP Publishing Ltd, Nov. 2019.

[42] O. Nikisins, A. George, and S. Marcel, "Domain adaptation in multi-channel autoencoder based features for robust face anti-spoofing," in *International Conference on Biometrics 2019, IEEE*, 2019.

[43] R. Oliveira, J. Kämpf, R. Vicente, R. Almeida, and A. Figueiredo, "Co2 experimental measurements towards the development of a predictive framework using user actions in smart buildings," in *Journal of Physics: Conference Series*, vol. 1343, IOP Publishing Ltd, Nov. 2019.

[44] N. Pappas and J. Henderson, "Deep residual output layers for neural language generation," in *Proceedings of the 36th International Conference on Machine Learning (ICML)*, 2019.

[45] S. Parida, "Overview of the 6th workshop on asian translation," in *Proceedings of the 6th Workshop on Asian Translation*, Hong Kong, China: Association for Computational Linguistics, Nov. 2019, 1–35.

[46] S. Parida and P. Motlicek, "Abstract text summarization: A low resource challenge," in *In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2019)*, HongKong, China: Association for Computational Linguistics (ACL), Nov. 2019, p. 5.

[47] S. Parida and P. Motlicek, "Idiap abstract text summarization system for german text summarization task," in *Proceedings of the 4th edition of the Swiss Text Analytics Conference*, Oct. 2019.

[48] S. Parida and P. Motlicek, "Idiap nmt system for wat 2019 multimodal translation task," in *Proceedings of the 6th Workshop on Asian Translation*, Hong Kong, China: Association for Computational Linguistics, Nov. 2019, 175–180.

[49] T.-T. Phan, S. Muralidhar, and D. Gatica-Perez, "Drink or drunk: Multimodal signals and drinking practices on instagram," in *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare*, Trento, Italy, May 2019.

[50] K. H. Poon and J. Kämpf, "A morphological based pv generation and energy consumption predictive model for singapore neighbourhood," in *Journal of Physics: Conference Series*, vol. 1343, IOP Publishing Ltd, Nov. 2019.

[51] D. N. Popa, J. Perez, J. Henderson, and E. Gaussier, "Implicit discourse relation classification with syntax-aware contextualized word representations," in *Proceedings of the 32nd International Florida Artificial Intelligence Research Society Conference*, 2019.

[52] D. N. Popa, J. Perez, J. Henderson, and E. Gaussier, "Satoke: How can syntax-aware contextualized word representations benefit implicit discourse relation classification?" In *Ptroc. 2019 Conférence sur l'Apprentissage automatique*, 2019.

[53] R. Raghavendra, S. Venkatesh, K. B. Raja, S. Bhattacharjee, P. Wasnik, S. Marcel, and C. Busch, "Custom silicone face masks - vulnerability of commercial face recognition systems & presentation attack detection," in *Proceedings of 7th IAPR/IEEE International Workshop on Biometrics and Forensics*, May 2019.

[54] D. Ram, L. Miculicich, and H. Bourlard, "Multilingual bottleneck features for query by example spoken term detection," in *IEEE Automatic Speech Recognition and Understanding Workshop*, 2019.

[55] N. Sacchi, A. Nanchen, M. Jaggi, and M. Cernak, "Open-vocabulary keyword spotting with audio and text embeddings," in *Proceedings of 2019 IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2019.

[56] B. Schnell and P. N. Garner, "Neural vtln for speaker adaptation in tts," in *Proc. 10th ISCA Speech Synthesis Workshop*, ISCA, Vienna, Austria, Sep. 2019, p. 6.

[57] R. Siegfried, Y. Yu, and J.-M. Odobez, "A deep learning approach for robust head pose independent eye movements recognition from videos," in *2019 ACM Symposium on Eye Tracking Research & Applications*, ACM, Jun. 2019, p. 5.

[58] J. Silverio, Y. Huang, F. J. Abu-Dakka, L. Rozo, and D. G. Caldwell, "Uncertainty-aware imitation learning using kernelized movement primitives," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2019.

[59] S. Srinivas and F. Fleuret, "Full-gradient representation for neural network visualization," in *Advances in Neural Information Processing Systems*, 2019.

[60] M. Suomalainen, S. Calinon, E. Pignat, and V. Kyrki, "Improving dual-arm assembly by master-slave compliance," in *Proc. IEEE Intl Conf. on Robotics and Automation*, 2019, pp. 8676–8682.

[61] M. Tammen, I. Kodrasi, and S. Doclo, "Joint estimation of retf vector and power spectral densities for speech enhancement based on alternating least squares," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, May 2019, pp. 795–799.

[62] L. Tarantino, P. N. Garner, and A. Lazaridis, "Self-attention for speech emotion recognition," in *Proc. Interspeech 2019*, 2019.

[63] C. Thomas, D. Jayagopi, and D. Gatica-Perez, "Booktubing across regions: Examining differences based on nonverbal and verbal cues," in *Proc. ACM Int. Conf. on Interactive Experiences for Television and Online Video (TVX)*, Salford, ENGLAND, Jun. 2019.

[64] S. Tong, P. N. Garner, and H. Bourlard, "An investigation of multilingual asr using end-to-end lf-mmi," in *International Conference on Acoustics, Speech and Signal Processing*, 2019.

[65] S. Tong, A. Vyas, P. N. Garner, and H. Bourlard, "Unbiased semi-supervised lf-mmi training using dropout," in *Proceedings of Interspeech 2019*, 2019.

[66] S. Tornay, M. Razavi, N. C. Camgoz, R. Bowden, and M. Magimai.-Doss, "Hmm-based approaches to model multichannel information in sign language inspired from articulatory features-based speech processing," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2019.

[67] S. Tulyakov, F. Fleuret, M. Kiefel, P. Gehler, and M. Hirsch, "Learning an event sequence embedding for event-based deep stereo," in *Proceedings of the IEEE International Conference on Computer Vision*, 2019.

[68] J. Vázquez-Canteli, T. Detjeen, G. Henze, J. Kämpf, and Z. Nagy, "Multi-agent reinforcement learning for adaptive demand response in smart cities," in *Journal of Physics: Conference Series*, vol. 1343, IOP Publishing Ltd, Nov. 2019.

[69] J. Vázquez-Canteli, J. Kämpf, G. Henze, and Z. Nagy, "Citylearn v1.0: An openai gym environment for demand response with deep reinforcement learning," in *Proceedings of the 6th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation*, New-York, USA: ACM, Nov. 2019, pp. 356–357.

[70] A. Vyas, P. Dighe, S. Tong, and H. Bourlard, "Analyzing uncertainties in speech recognition using dropout," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Jul. 2019.

[71] H. Wang, J. Henderson, and P. Merlo, "Weakly-supervised concept-based adversarial learning for cross-lingual word embeddings," in *Proc. 2019 Conference on Empirical Methods in Natural Language Processing*, 2019.

[72] Y. Wu, J. Kämpf, and J. L. Scartezzini, "Daylight regulated by automated external venetian blinds based on hdr sky luminance mapping in winter," in *Journal of Physics: Conference Series*, vol. 1343, IOP Publishing Ltd, Nov. 2019.

[73] Y. Yu, G. Liu, and J.-M. Odobez, "Improving few-shot user-specific gaze adaptation via gaze redirection synthesis," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019.

[74] Q. Zhan, P. Motlicek, S. Du, Y. Shan, X. Xie, and S. Ma, "Cross-lingual automatic speech recognition exploiting articulatory features," in *Proceedings of APSIPA ASC 2019*, 2019.

6 List of Patents

While evaluating the success of technology transfer is a difficult task, invention disclosures and patents are usually key metrics in the field. As a necessary first step to any patent filing, an invention disclosure aims to identify a technology with a certain level of maturity and some promising economic potential.

In 2019, 27 invention and software disclosures have been filled. A patent committee examines each invention disclosure and decides to move forward with a patent filing or not. The committee is composed of the head of technology transfer, two representatives of the direction and one representative of the researchers. Since 2012, 10 patents have been filed: 3 patents are granted, 6 applications are pending, one application has been sold and 2 applications are not published yet. Following are listed the granted patents and the published pending applications.

6.1 Granted patents

US 9,689,959 B2 A. Asaei, H. Bourlard, V. Cevher, "Method, apparatus and computer program product for determining the location of a plurality of speech sources"

US 9,058,541 B2 C. Dubout, F. Fleuret, "Object detection method, object detector and object detection computer program"

US 9,973,503 B2 S. Marcel, A. Anjos, P. Abbet, "Method and internet-connected server for reviewing a computer-executable experiment"

6.2 Pending applications

US 2014/0149104 A1, EP 2736042 A1 N-H. Kim, P. Motlicek, P. N. Garner, D. Imseng, J-W. Lee, J-M. Cho, "Apparatus and method for constructing multilingual acoustic model and computer readable recording medium for storing program for performing the method"

WO 2015/192879 A1 K. A. Funes Mora, J-M. Odobez, "A gaze estimation method and apparatus"

US 2017/0069306 A1 A. Asaei, M. Cernak, H. Bourlard, "Signal processing method and apparatus based on structured sparsity of phonological features"

WO 2017/221049 A1, US 2019/0180040 A1 A. Anjos, S. Marcel, "A data-network connected server, a device, a platform and a method for conducting computer-executable experiments"