
RAPPORT ANNUEL 2017



La page de couverture montre la plateforme MMD: une plateforme de recherche reproductible testée à l'Idiap pour détecter des maladies de plantes, en apprenant par des démonstrations comment manipuler et inspecter les feuilles.

réalisation Martina Fellay, Léonore Miauton, équipe gestion de projets

rédaction Laurence Facelli, Marie Rumignani, Idiap Project Management Team

correction Marie-Jeanne Krill **traduction** Dave Brooks, ELCS.ch

crédits photographiques Céline Ribordy, Sion · Sedrik Nemeth, Sion · Vincent Spano · Idiap, Martigny

conception graphique Formaz — Andenmatten, Sion **typographies** Cera · Franziska **papier** Planojet® FSC

impression ronquozgraphix.ch, Sion **tirage** 1'500 exemplaires

PRINCIPAUX PARTENAIRES

VILLE DE MARTIGNY

CANTON DU VALAIS

CONFÉDÉRATION SUISSE

Secrétariat d'Etat à la formation, à la recherche et à l'innovation (SEFRI)



www.loro.ch



www.swisscom.ch



www.epfl.ch



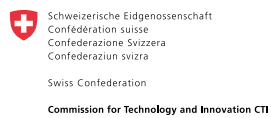
www.theark.ch



www.ideark.ch



www.snf.ch



www.kti.admin.ch



www.haslerstiftung.ch

Messages

- 04** L'Intelligence Artificielle ou l'Idiap à la pointe de la recherche
- 05** Qu'est-ce qui fait le succès de l'Idiap ?

Organisation

- 07** Idiap, carte d'identité
- 10** Domaines d'application
- 11** Organigramme opérationnel
- 12** Groupes de recherche
- 14** Conseil de fondation
- 15** Comité d'accompagnement
- 15** Collaborateurs

Finances

- 19** Compte d'exploitation
- 20** Sources de financement
- 21** Bilan

Recherche

- 23** MEMMO - Une mémoire du mouvement pour qu'un robot puisse s'adapter rapidement
- 24** Martigny au cœur d'un réseau européen d'étude sur le langage
- 25** L'Idiap est dans le move
- 26** L'intelligence Artificielle sous le feu des projecteurs de la recherche

Réseau

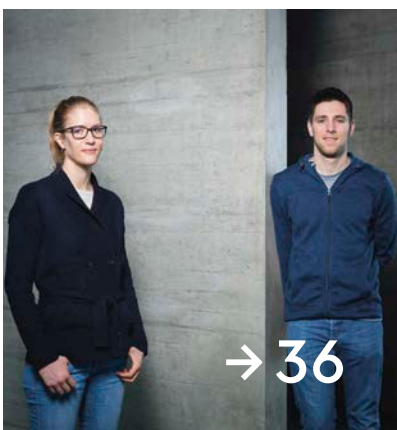
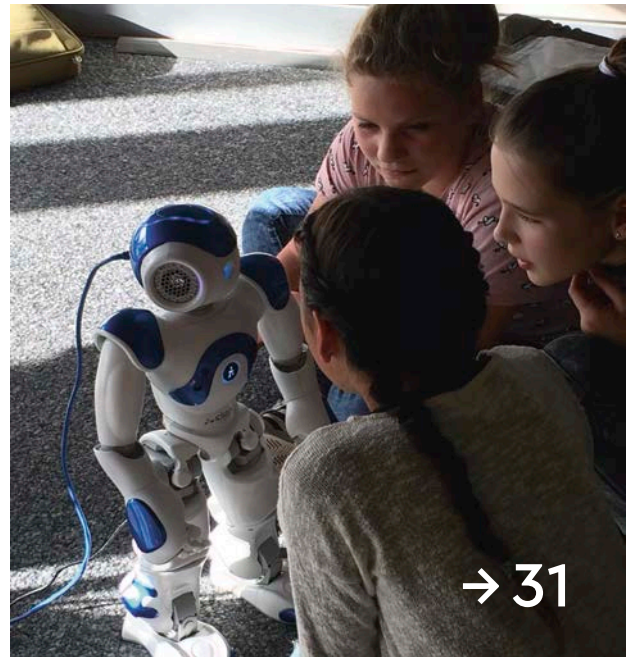
- 29** La 6^e édition de l'International Create Challenge
- 30** Se tourner vers le grand public, une vocation pour l'Idiap
- 31** Activités pour le jeune public
- 32** Un rayonnement sans frontière pour les spin-off de l'Idiap

Visages

- 35** Du MIT à Martigny, l'odyssée d'un chercheur globe-trotteur passionné de langage
- 36** Mélanie Huck et Guillaume Clivaz fêtent leur premier anniversaire au sein de l'Idiap
- 37** Marie-Constance Kaiflin, une juriste au carrefour de l'Intelligence Artificielle
- 38** Distinctions
- 39** Thèses achevées

Encarté scientifique

- II** Scientific Report
- XXVII** Main projects in progress
- XXXII** Major publications / Conferences



L'Intelligence Artificielle ou l'Idiap à la pointe de la recherche

L'exercice 2017 est marqué par la progression du chiffre d'affaires de l'Institut dû principalement à la diversité des sources de revenus issus notamment de projets nord-américains et à la notoriété grandissante du Centre en biométrie. L'engagement comme l'excellence de notre direction et des collaborateurs doivent également être relevés.

Nous pouvons compter depuis longtemps sur des subventions publiques et remercions les autorités politiques et les membres des administrations qui autorisent ainsi le déploiement d'activités de recherche fondamentale qui nous permettent l'obtention de mandats, notamment dans l'industrie.

L'évolution de l'Institut prouve que les domaines d'activités dans lesquels il œuvre sont en pleine expansion. Nul ne peut contester les opportunités que l'Intelligence Artificielle offre désormais, surtout si l'on intègre la capacité des machines à développer une forme de pensée leur permettant d'acquérir une réelle autonomie. Les derniers robots industriels et les applications en médecine représentent les précurseurs d'une ère dans laquelle l'entier du secteur tertiaire s'immisce d'ores et déjà.

Ce n'est pas sans fierté que l'Idiap participe à cette évolution. Quelques conditions sont toutefois requises pour surfer sur la vague du succès. Nous pensons ici à la nécessité impérieuse de disposer d'ingénieurs de développement informatique, capables de « mettre en musique » les résultats des recherches entreprises au sein de l'Idiap. L'attractivité de ce métier doit être renforcée à dessein de pérenniser nos activités. Un programme de formation spécifique offert à de jeunes talents sera prochainement mis sur pied en interne.

En plus de la recherche fondamentale, l'Idiap assure toujours une activité importante dans le transfert de



« Nous sommes convaincus de la pertinence de notre activité en Intelligence Artificielle »

Olivier Dumas

Président du Conseil de fondation de l'Idiap

technologie. Nombre de start-up ont prospéré en son sein. Une seconde voie s'ouvre désormais : celle d'unités d'application de technologies dans lesquelles l'Idiap évolue. Après la création en 2014 du Centre suisse en biométrie, notre Institut collabore maintenant avec le CREM (Centre de recherches énergétiques et municipales) afin de mettre en relation les connaissances des deux entités. Il offre de plus ses services et compétences dans plusieurs domaines : ressources humaines, communication, gestion de projet ou encore services comptables et financiers.

Plus que jamais, nous sommes convaincus de la pertinence de notre activité dans la communauté de l'Intelligence Artificielle. Alors que certains voient là les prémices d'un asservissement de l'humanité, nous sommes persuadés qu'aucune menace spécifiquement liée à l'Intelligence Artificielle n'existe. Seul son emploi abusif représente un danger bien réel sans oublier qu'à chaque fois que la société acquiert une technologie de rang supérieur, la science est incriminée de mille maux. Nous osons croire en la sagesse de l'homme pour, une fois encore, ne pas tomber dans de fâcheux travers.

Qu'est-ce qui fait le succès de l'Idiap ?

En réponse à l'invitation d'un grand groupe industriel européen, je suis récemment allé présenter l'Institut de Recherche Idiap. J'étais alors heureux d'apprendre que leurs larges besoins en Intelligence Artificielle (IA) se retrouvaient parfaitement dans nos offres de R&D.

Cette observation s'explique par le fait que l'Idiap est aujourd'hui le résultat d'une construction aux bases solides. Chaque brique qui la compose provient du travail de nos chercheurs, assistants et doctorants, patiemment répertorié depuis plus de 25 ans. Des efforts sont constamment déployés afin d'encourager nos nouveaux arrivés à développer leurs activités sur ces bases solides. Grâce à sa longue expérience, et une bonne exploitation des connaissances acquises, aussi bien au niveau de la recherche fondamentale que du transfert de technologie vers l'industrie, l'Idiap est devenu une référence dans le domaine de l'IA au service de la société. Etant à l'origine d'un grand nombre de logiciels et de bases de données utilisés partout dans le monde, notre Institut fait preuve d'une excellente maîtrise de ses ressources.

Au niveau de la recherche fondamentale en IA, il n'y a pas eu dernièrement de grandes percées comme celles survenues dans les années 80 et 90. Par contre, les opportunités et le renouveau du domaine résident dans l'exploitation des capacités de stockage et de calcul toujours plus grandes. À cela s'ajoute une quantité quasi illimitée de données très variées à disposition, le Big Data. Et voilà que ce qui avait été imaginé il y a plus de vingt ans devient doucement réalité, tout en soulevant de nouveaux challenges, aussi bien scientifiques qu'éthiques.

La force de l'Institut est la cohérence et l'indépendance de ses groupes de recherche. Actuellement au nombre de neuf, ils couvrent la majorité des domaines clés en Intelligence Artificielle, résultant en une synergie particulièrement productive et compétitive lorsqu'ils travaillent ensemble sur des projets communs.



« L'Idiap, un *one-stop shop* en Intelligence Artificielle »

Hervé Boulard

Directeur de l'Idiap, professeur EPFL

Dans ce contexte, le choix d'un nouveau groupe est toujours un pari risqué, ou une idée folle. L'Idiap ne veut pas simplement suivre les modes. Pour garder son excellence, il faut se diversifier autour de nos domaines privilégiés, tout en anticipant les besoins futurs.

La création en 2015 du groupe «Uncertainty Quantification and Optimal Design» illustre ce propos. Alors que la tendance actuelle exploite surtout le Big Data, ce groupe de recherche s'intéresse à résoudre des problèmes n'ayant à disposition qu'une faible quantité de données (small data). Il s'agit alors de développer un modèle à partir d'une faible quantité d'échantillons. Le savoir développé dans ce cadre est potentiellement applicable à un grand nombre de cas, par exemple dans le placement optimal de senseurs en vue d'évaluer des risques environnementaux.

Voilà comment l'Idiap, en s'appuyant sur ses connaissances approfondies, le dynamisme de chacun de ses groupes de recherche, son équipe de R&D garantissant la qualité de ses logiciels et leur transfert vers l'industrie, propose une offre unique dans le domaine de l'IA.



organisation

06 — 17

Un institut à taille humaine et au rayonnement international

Fort d'une centaine de collaborateurs et de domaines de recherche en lien avec les défis actuels, l'Idiap s'engage pour un progrès scientifique au service de l'homme.

En 1991 déjà, date de sa fondation, l'Institut de Recherche Idiap se donnait pour mission de s'engager pour un progrès scientifique au service du bien-être des hommes. Aujourd'hui, vingt-six ans plus tard, l'institut place toujours les intérêts de la société au cœur de sa démarche.

Déferlement d'outils technologiques

En ce début de XXI^e siècle, on assiste à un déferlement permanent de nouveaux outils technologiques. S'ils permettent des gains considérables en termes de productivité et de confort, ils bouleversent aussi les habitudes des gens, laissant une part des utilisateurs démunis et une autre lassée par la modification récurrente des systèmes. Dans ce contexte, l'Idiap – souhaitant mettre l'intelligence artificielle au service de la

société - travaille essentiellement à l'amélioration des relations homme-machine, et à l'optimisation de la communication humaine.

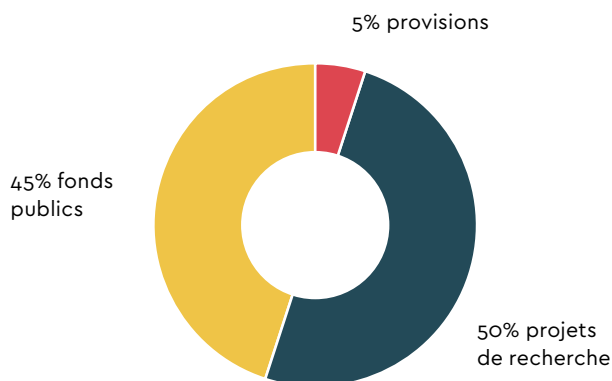
Réseau national et international

Au niveau suisse, l'Idiap travaille avec les écoles polytechniques, les hautes écoles et les universités où se pratiquent également des activités de recherche. Impliqué dans plusieurs projets européens, l'institut entretient des liens étroits avec de nombreux partenaires, essentiellement en France, en Angleterre et en Allemagne. Outre-Atlantique, c'est principalement en Californie, avec entre autre l'International Computer Science Institute (ICSI) de Berkeley et l'Information Sciences Institute (ISI, USC Viterbi), que l'on retrouve les partenaires privilégiés.

Structure

- Fondation à but non lucratif
- Institution indépendante (mais affiliée à l'EPFL par un plan de développement commun)
- Spécialisée dans l'Intelligence Artificielle, la gestion de l'information multimédia et l'interaction multimodale homme-machine
- Fondée en 1991 par :
 - la Ville de Martigny
 - l'Etat du Valais
 - l'EPFL
 - l'Université de Genève
 - Swisscom

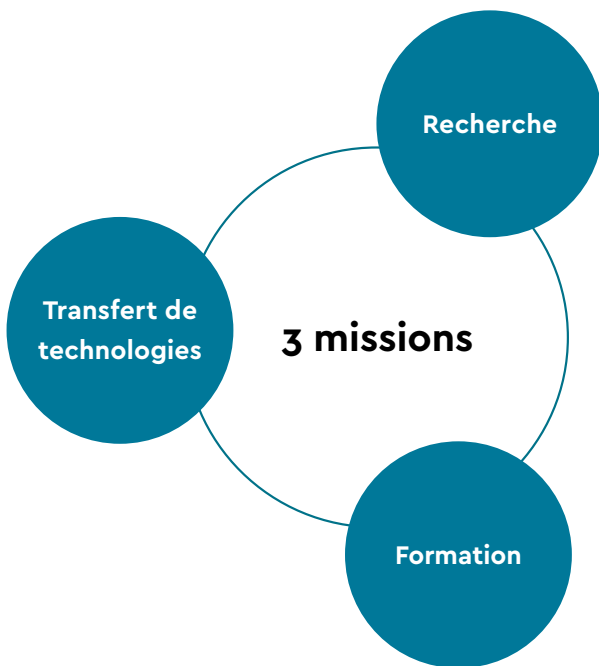
Financement



Les 3 missions de l'Idiap

L'Institut de Recherche Idiap est une institution reconnue par la Confédération comme faisant partie du domaine stratégique des EPF et soutenue en vertu de la «Loi fédérale sur l'encouragement de la recherche et l'innovation» (art.15). Sa vision est de promouvoir la qualité de la vie par

le progrès scientifique dans le domaine de l'Intelligence Artificielle, notamment en lien avec le traitement de l'information multimédia et sensorielle. Les activités de l'Idiap s'orientent selon 3 missions complémentaires :



8

Recherche

Mener des projets de recherche fondamentale au plus haut niveau dans ses domaines de prédilection, s'assurant ainsi une place parmi les meilleurs à l'échelle nationale, européenne et mondiale. Sur la scène internationale, l'Idiap bénéficie d'un large réseau de partenaires et collabore activement avec de grandes universités, des centres de recherche publics ou privés, etc.

Les projets de recherche compétitifs garantissent 50% du financement de l'institut.

Formation

Former la relève en faisant découvrir le monde de la recherche à des stagiaires, en accueillant de jeunes chercheurs talentueux préparant leur doctorat, et en dispensant de nombreux cours à l'EPFL et en interne.

Un collaborateur scientifique sur deux est un doctorant.

Transfert de technologies

Assurer le transfert des technologies à travers la dissémination la plus large possible de ses résultats de recherche dans la communauté scientifique, mais aussi et surtout en tissant des liens étroits avec le monde industriel.

Le transfert de technologies représente 10% du budget de l'Idiap.

Grâce à l'incubateur The Ark sur le site d'IdeArk, l'Idiap permet l'éclosion de nombreuses start-up.

L'Idiap en chiffres 2017

Ressources humaines

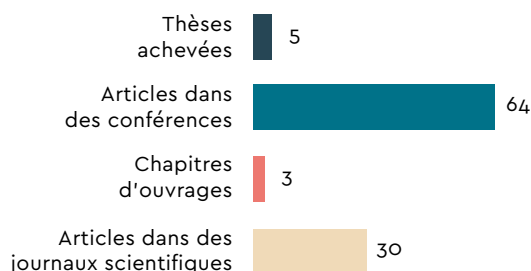
- 4 professeurs
 - 2 maîtres d'enseignement et de recherche (MER)
 - 6 chercheurs permanents
 - 7 chercheurs associés
 - 18 postdoctorants
 - 33 doctorants
 - 11 ingénieurs de développement
 - 7 ingénieurs système
 - 7 collaborateurs administratifs
 - 13 stagiaires (moyenne/année)
- 144 personnes au total**
98 équivalents plein temps

Activités scientifiques

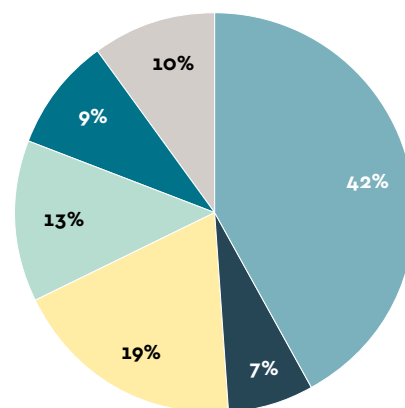
- Participation en 2017 à **65 programmes de recherche**
- Direction de projet dans plus de **45 consortiums** (Suisse, Europe, projets industriels)
- Participation à la stratégie de développement économique du Canton du Valais à travers le programme The Ark et en particulier la société IdeArk
- **Plus de 50 postes** dans les start-up IdeArk

Publications

En 2017, les scientifiques de l'Idiap ont participé à 102 publications revues par les pairs.



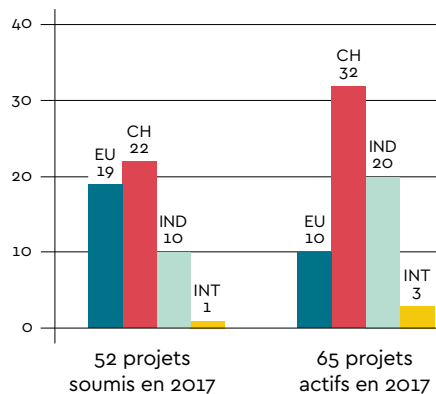
Au total, 25 nationalités sont représentées au sein de l'Idiap



- Suisse
- Amérique
- France
- Reste Europe
- Inde
- Reste Asie
- Apatrides, 4 robots

Soumission et financement des projets de recherche

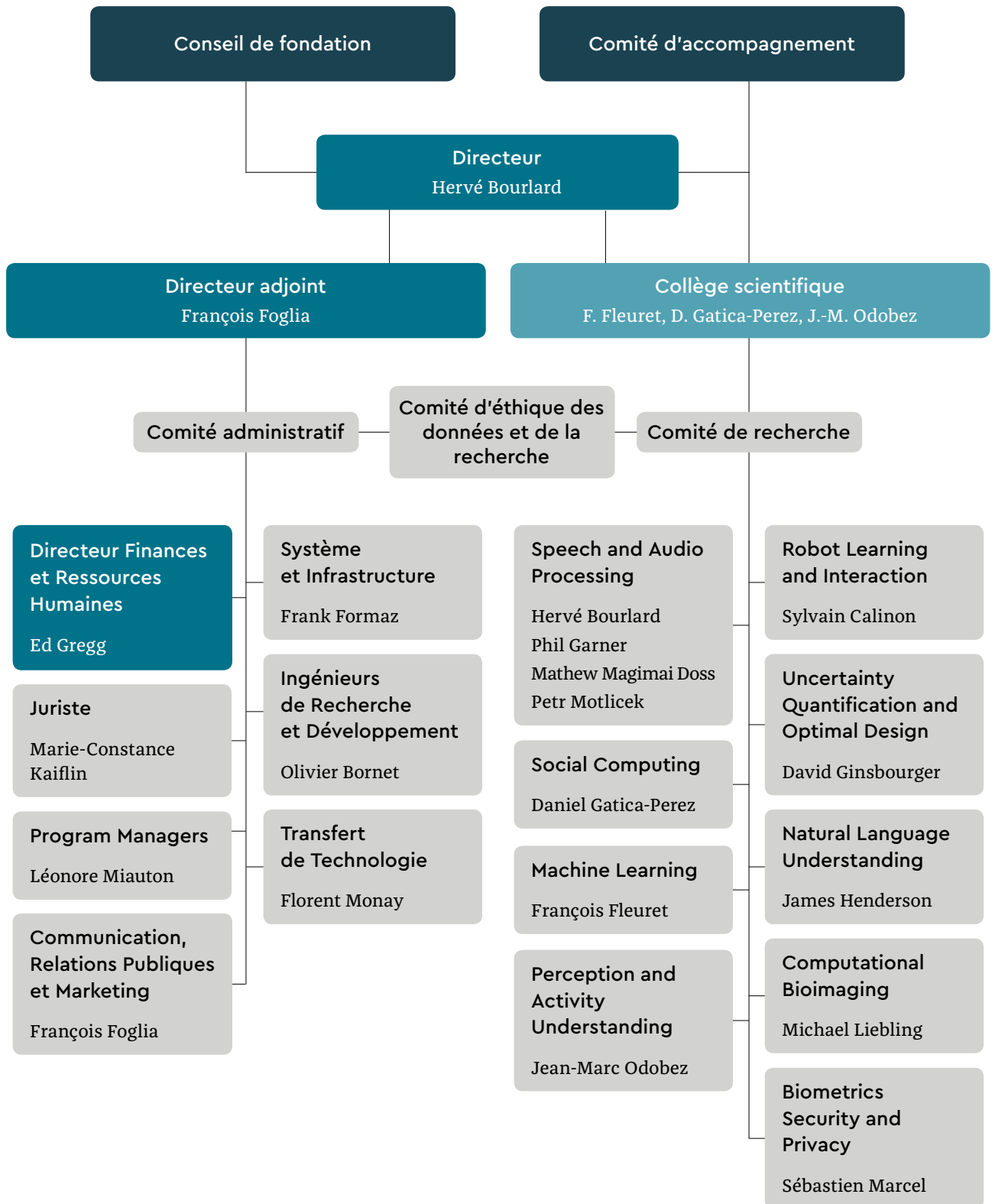
En 2017, l'Idiap a soumis 52 projets (22 à une agence de financement suisse, 19 à un fonds européen, 10 projets industriels et 1 projet international). En parallèle, l'institut a géré 65 projets actifs durant l'année. Dans les deux cas, c'est en Suisse que l'activité a été la plus intense.



5 thèmes de recherche 10 domaines d'application



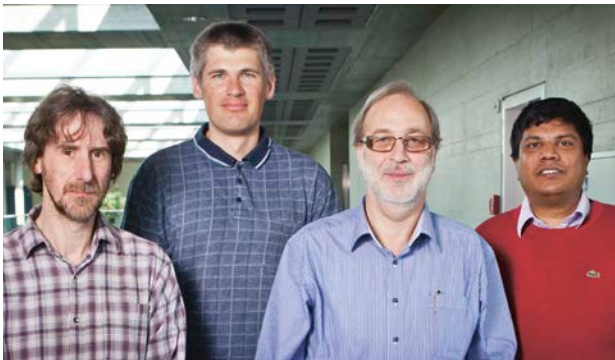
Organigramme opérationnel



Groupes de recherche

Speech & Audio Processing

P^r Hervé Bourlard (3^e depuis la gauche),
D^r Phil Garner, D^r Petr Motlicek, D^r Mathew Magimai-Doss
H-index : PG : 23, PM : 16, HB : 59, MMD : 25



À l'Idiap, les travaux sont menés au sein de neuf groupes de recherche. Celui chargé du traitement de la parole et du son est emmené par le directeur de l'institut, Hervé Bourlard, et trois chercheurs permanents. Les autres groupes sont également pilotés par des chercheurs permanents dont un professeur titulaire et deux maîtres d'enseignement et de recherche (MER).

Le traitement de la parole est depuis de nombreuses années l'un des thèmes de recherche majeurs de l'Idiap. L'institut occupe d'ailleurs une place de choix sur la scène internationale dans ce domaine. Au cœur des travaux du groupe de recherche le plus important de l'Idiap : la reconnaissance automatique de la parole de manière statistique, la synthèse vocale et le traitement générique de l'information audio (localisation de la source, des réseaux de microphones, segmentation par locuteur, indexation de l'information, codage du signal vocal à de très bas débits, analyse du bruit de fond).

12



Social Computing

P^r Daniel Gatica-Perez
H-index : 55

Le « Social Computing » est un domaine interdisciplinaire qui intègre les théories et les modèles de l'informatique mobile et omniprésente, de l'apprentissage automatique, du multimédia et des sciences sociales. Tous ces domaines sont réunis pour détecter, analyser et interpréter les comportements humains et sociaux au quotidien dans le but de créer des instruments qui soutiennent les interactions et la communication. Les lignes de recherche actuelles comprennent l'étude des interactions face-à-face, l'analyse comportementale de vidéos sociales, le crowdsourcing et le traitement de données urbaines à l'aide de smartphones et de réseaux sociaux mobiles.



Computer Vision & Learning

D^r François Fleuret
H-index : 29

Ce groupe s'intéresse au développement de nouvelles techniques d'apprentissage statistique, essentiellement pour la vision par ordinateur, avec un intérêt particulier pour leurs propriétés computationnelles. Les principales applications sur lesquelles le groupe travaille sont l'analyse automatique d'images, en particulier l'extraction automatique de sémantique, la détection d'objets, et le suivi de personnes et de structures biologiques.

* **Le H-index** tente de quantifier la productivité et l'impact d'un scientifique en fonction du niveau de citation de ses publications. Plus le chiffre est élevé, plus le niveau de citation est important.



Perception & Activity Understanding

Dr Jean-Marc Odobez
H-index : 40

Ce groupe de recherche s'intéresse à l'analyse des activités humaines à partir de données multimodales. Cette analyse repose sur le développement d'algorithmes fondés sur des méthodes de vision par ordinateur, d'apprentissage, et de fusion de données pour résoudre des tâches fondamentales comme la détection et le suivi d'objets et de personnes, leur représentation et la caractérisation de leur état, ainsi que la modélisation de données séquentielles et leur interprétation sous forme de gestes, comportements ou relations sociales. La surveillance, l'analyse de comportements, les interfaces homme-robot et l'analyse de contenus multi-médias constituent les principaux domaines d'application.



Robot Learning & Interaction

Dr Sylvain Calinon
H-index : 35

Le groupe Apprentissage & Interaction robotiques cible des applications robotiques centrées sur l'humain. L'objectif scientifique est de développer des approches statistiques pour encoder les mouvements et comportements de robots évoluant dans des environnements non contraints. Dans ces applications, les modèles ont des rôles multiples (reconnaissance, prédiction, reproduction) et sont partagés par des stratégies d'apprentissage diverses (imitation, émulation, correction incrémentale ou exploration). Le but est de faciliter le transfert des tâches de l'utilisateur au robot, ou entre robots, en exploitant des informations sensorielles multimodales et en développant des interfaces d'apprentissage intuitives.



Uncertainty Quantification and Optimal Design

Pr David Ginsbourger
H-index : 22

Ce groupe de recherche s'intéresse à la quantification et la réduction d'incertitudes dans le contexte de modèles de haute fidélité, avec une expertise principale sur les méthodes de processus gaussiens et la planification d'expériences numériques pour l'optimisation, l'inversion, et autres problèmes apparentés. Les domaines d'application incluent notamment l'énergie et les géosciences, avec des collaborations allant de l'ingénierie de sûreté à l'hydrologie en passant par les sciences du climat.



Computational Bio Imaging Group

Pr Michael Liebling
H-index : 21

La recherche dans ce groupe se concentre sur l'imagerie computationnelle et l'analyse d'images biomédicales. Ceci inclut le développement d'algorithmes pour la déconvolution et la super-résolution en microscopie optique ainsi que la reconstruction tomographique tridimensionnelle et, plus généralement, la combinaison de dispositifs et procédés de détection inhabituels avec des logiciels de calcul pour produire des images idéalement adaptées à l'observation et la quantification de systèmes biologiques complexes et vivants.



Biometrics Security and Privacy

Dr Sébastien Marcel
H-index : 43

En informatique, la biométrie se rapporte à la reconnaissance automatique d'individus par l'utilisation de caractéristiques comportementales et biologiques. Les chercheurs du groupe Biométrie étudient et développent de nouveaux algorithmes de traitement d'image et de reconnaissance des formes pour la reconnaissance de visage (2D, 3D et proche infrarouge), la reconnaissance du locuteur, la détection d'attaques (antispoofing) ainsi que pour certaines modalités biométriques émergentes (électrophysiologie et veines). Le groupe encourage la reproduction des résultats de recherche et le transfert de technologies en utilisant sa propre librairie de traitement du signal et d'apprentissage automatique par ordinateur.



Natural Language Understanding

Dr James Henderson
H-index : 22

Le groupe Compréhension des langues a été créé en septembre 2017 sous la direction de James Henderson, en partie en tant que continuation du précédent groupe Traitement automatique des langues d'Andrei Popescu-Belis. Le groupe continue le travail sur l'analyse des textes au niveau sémantique et pragmatique pour améliorer les performances de la traduction automatique et de la recherche d'informations, en mettant récemment l'accent sur la traduction automatique neurale et sur des modèles de deep learning basés sur l'attention. Cela s'intègre bien à la nouvelle direction de recherche du groupe qui se tourne vers la prédiction structurée des réseaux neurologiques et l'apprentissage représentatif pour la modélisation sémantique et syntaxique des textes et discours, comprenant la modélisation de l'abstraction (implication textuelle) et la synthétisation.

Conseil de fondation

Le Conseil de fondation assume la responsabilité de gestion économique et financière de l'institut de recherche, définit ses structures, nomme son directeur, et de manière plus générale veille au bon développement de la fondation en défendant ses intérêts.



M. Olivier Dumas, président

Administrateur indépendant
et conseiller d'entreprises



M. Jordi Montserrat

Directeur régional de
Venturelab



**M^{me} Anne-Laure Couchepin
Vouilloz, vice-Présidente**

Présidente de la ville de
Martigny



Pr Stéphane Marchand-Maillet

Professeur associé à la Faculté
des sciences de l'Université de
Genève



M. Marc-André Berclaz

Directeur opérationnel de
l'Antenne EPFL Valais-Wallis



M. Dominique Perruchoud

Président du Conseil d'adminis-
tration de Cimark



M. Stefan Bumann

Chef de service des Hautes
Ecoles (SHE)



Pr Pierre Vandergheynst

Professeur et vice-président
pour l'Education de l'EPFL



M. Patrick Furrer

Collaborateur scientifique
chez swissuniversities



M. David Watrin

Responsable de l'unité
de produit « Sécurité &
Intelligence » chez Swisscom

Conseil stratégique international

Le comité d'accompagnement scientifique se compose de personnalités du monde scientifique choisies par la Direction de l'Idiap pour leurs compétences exceptionnelles et leur vision avant-gardiste. Bien que leur rôle soit strictement consultatif, leur soutien et leurs conseils sont fréquemment sollicités et se révèlent précieux dans la prise de décision en matière de recherche, de formation et de transfert de technologies.



Dr Alex Acero

Senior Director at Apple, Cupertino, CA, USA



Prof Klaus-Robert Müller

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



Dr D. Vanessa Evers

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



Dr Prem Natarajan

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



Prof Anil K. Jain

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



Prof Bernt Schiele

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



Prof Johanna Moore

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



Dr Luciana Vaccaro

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

Collaborateurs

Personnel scientifique

Prénom, nom, fonction, origine, date d'entrée en fonction

16 Afsaneh Asaei, postdoctorante, Iran, 2008
 Sushil Bhattacharjee, research associate, Suisse, 2016
 Hervé Bourlard, directeur, Belgique/Suisse, 1996
 Sylvain Calinon, chercheur, Suisse, 2014
 Yuanzhouhan Cao, postdoctorant, Chine, 2017
 Gülcan Can, doctorante, Turquie, 2013
 Olivier Canévet, doctorant, France, 2012
 Tatjana Chavdarova, doctorante, Macédoine, 2014
 Tiago De Freitas Pereira, doctorant, Brésil, 2014
 Subhadeep Dey, doctorant, Inde, 2014
 Pranay Dighe, doctorant, Inde, 2013
 Pavankumar Dubagunta, doctorant, Inde, 2017
 François Fleuret, chercheur senior, France, 2007
 Philip Garner, chercheur senior, Angleterre, 2007
 Daniel Gatica-Perez, chercheur senior, Mexique, 2002
 David Geissbuhler, postdoctorant, Suisse, 2017
 Anjith George, postdoctorant, Inde, 2017
 David Ginsbourger, chercheur senior, France, 2015
 Weipeng He, doctorant, Chine, 2016
 James Henderson, chercheur senior, Etats-Unis 2017
 Guillaume Heusch, collaborateur scientifique, Suisse, 2015
 Rui Hu, postdoctorante, Chine, 2013
 David Imseng, postdoctorant, Suisse, 2009
 Christian Jaques, doctorant, Suisse, 2016
 Noémie Jaquier, doctorante, Suisse, 2016
 Cijo Jose, doctorant, Inde, 2013
 Angelos Katharopoulos, doctorant, Grèce, 2017
 Ina Kodrasi, postdoctorante, Albanie, 2017
 Alain Komaty, postdoctorant, France, 2016
 Pavel Korshunov, postdoctorant, Estonie, 2015
 Vedrana Krivokuca, postdoctorante, Nouvelle-Zélande, 2017
 Thibaut Kulak, doctorant, France, 2017
 Andras Kupcsik, postdoctorant, Hongrie, 2017
 Florian Labhart, doctorant, Suisse, 2017
 Nam Le, doctorant, Vietnam, 2015
 Yann Lepoittevin, postdoctorant, France, 2016
 Michael Liebling, chercheur senior, Suisse, 2014
 Gang Liu, postdoctorant, Chine, 2017
 Srikanth Madikeri, postdoctorant, Inde, 2013
 Mathew Magimai Doss, chercheur, Inde, 2007
 Sébastien Marcel, chercheur senior, France/Suisse, 2000
 Olivia Mariani, doctorante, Suisse, 2016

Angel Martinez Gonzalez, doctorant, Mexique, 2016
 Amir Mohammadi, doctorant, Iran, 2016
 Petr Motlicek, chercheur, République tchèque, 2005
 Hannah Muckenhirn, doctorante, France, 2015
 Skanda Muralidhar, doctorant, Inde, 2014
 James Newling, doctorant, Angleterre, 2013
 Laurent Nguyen, postdoctorant, Suisse, 2011
 Olegs Nikisins, postdoctorant, Lettonie, 2017
 Jean-Marc Odobez, chercheur senior, France/Suisse, 2001
 Nikolaos Pappas, doctorant, Grèce, 2012
 Trung Phan, doctorant, Vietnam, 2015
 Emmanuel Pignat, doctorant, Suisse, 2016
 Xiao Pu, doctorante, Chine, 2014
 André Rabello Dos Anjos, collaborateur scientifique, Brésil, 2010
 Dhananjay Ram, doctorant, Inde, 2014
 Ramya Rasipuram, postdoctorante, Inde, 2010
 Marzieh Razavi, doctorante, Iran, 2013
 Bastian Schnell, doctorant, Allemagne, 2017
 Jilt Sebastian, doctorant, Inde, 2017
 Adrian Shajkofci, doctorant, Suisse, 2016
 Rémy Siegfried, doctorant, Suisse, 2017
 Suraj Srinivas, doctorant, Inde, 2017
 Ajay Srinivasamurthy, postdoctorant, Inde, 2017
 Ajay Tanwani, doctorant, Pakistan, 2015
 Sibong Tong, doctorant, Chine, 2016
 Sandrine Tornay, doctorante, Suisse, 2016
 Michael Villamizar, postdoctorant, Espagne, 2017
 Bogdan Vlasenko, postdoctorant, Allemagne 2017
 Yu Yu, doctorant, Chine, 2015

Ingénieurs de développement

Philip Abbet, ing. de développement senior, Suisse, 2006
 Olivier Bornet, responsable ing. de développement, Suisse, 2004
 Milos Cernak, ing. de développement senior, Slovaquie, 2011
 Guillaume Clivaz, ing. de développement, Suisse, 2017
 Samuel Gaist, ing. de développement, Suisse, 2013
 Mélanie Huck, ing. de développement, Suisse, 2017
 Salim Kayal, ing. de développement, Suisse, 2011
 Christine Marcel, ing. de développement, France/Suisse, 2007
 Florent Monay, officier de transfert de technologie, Suisse, 2008
 Alexandre Nanchen, ing. de développement senior, Suisse, 2008
 Flavio Tarsetti, ing. de développement senior, Suisse, 2008

Personnel administratif

Prénom, nom, fonction, origine, date d'entrée en fonction

Elisa Bovio, junior program manager et aide-comptable, Suisse 2015
Martina Fellay, program manager, Autriche, 2012
François Foglia, adjoint du directeur, Suisse, 2006
Edward-Lee Gregg, responsable financier, Etats-Unis, 2004
Marie-Constance Kaiflin, conseillère juridique, Suisse, 2017
Léonore Miauton, responsable program managers, Suisse, 2012
Sylvie Millius, assistante administrative, Suisse, 1996
Nadine Rousseau, assistante administrative, Belgique/Suisse, 1998

Ingénieurs système

Bastien Crettol, ingénieur système, Suisse, 2005
Norbert Crettol, ingénieur système, Suisse, 2002
Cédric Dufour, ingénieur système, Suisse, 2007
Frank Formaz, administrateur système, Suisse, 1998
Louis-Marie Plumel, ingénieur système, France, 2011
Vincent Spano, webmaster, Suisse, 2004
Laurent Tomas, administrateur système, Suisse, 2017

Stagiaires

Prénom, nom, origine, institution d'origine (pays)

Les stagiaires de l'Idiap passent généralement entre trois et dix mois dans l'institut de recherche. Certains sont étudiants à l'EPFL et effectuent ce stage dans le cadre de leur travail de diplôme. D'autres arrivent dans le cadre de programmes d'échange d'étudiants mis en place dans les projets européens auxquels participe l'Idiap.

Dylan Bourgeois, France, EPFL
Jaden Diefenbaugh, Etats-Unis, Oregon Univ., Corvallis, USA
Selen Kabil, Turquie, Middle East Technical Univ., Ankara, Turquie
Wei Ma, Inde, Nanjing Univ., Chine
Luis Emmanuel Medina Rios, Mexique, EPFL
Zohreh Mostaani, Iran, Ozgegin Univ., Instabul, Turquie
Amrutha Prasad, Inde, Karnataka, Inde
Alain Rossier, Suisse, EPFL
Florian Standaert, Suisse, EPFL



finances

18 — 21

Compte d'exploitation (CHF)

PRODUITS	2016	2017	%
Commune de Martigny	700'000	700'000	6.3
Etat du Valais	1'720'000	2'000'000	17.9
Confédération	2'423'700	2'418'500	21.6
Loterie Romande	87'000	90'000	0.8
Projets Fonds national suisse de la recherche scientifique	1'537'844	1'268'033	11.3
Projets Fondation Hasler	175'781	224'838	2
Projets européens	1'404'423	1'760'239	15.7
Projets The Ark	245'380	215'650	1.9
Projets CTI	638'467	1'007'604	9
Projets USA	—	296'733	2.6
Contribution EPFL	72'000	72'000	0.6
Financement industriel	414'669	147'015	1.3
Autres financements / Produits exceptionnels	738'592	490'277	4.4
Provisions	555'000	510'000	4.6
TOTAL DES PRODUITS	10'712'856	11'200'889	100

CHARGES	2016	2017	%
Frais de personnel	7'948'334	8'588'346	76.7
Formation et déplacements	427'372	371'307	3.4
Partenaires externes	102'292	229'278	2.1
Informatique: matériel et maintenance	253'410	355'319	3.2
Frais administratifs	166'113	158'719	1.5
Promotion et communication	45'769	18'147	0.2
Loyer et charges	709'229	671'754	6
Amortissement	363'090	482'636	4.3
Charges exceptionnelles	—	—	—
Variation provisions	685'000	290'896	2.6
TOTAL DES CHARGES	10'700'609	11'166'402	100
RÉSULTAT D'EXPLOITATION	12'247	34'487	—

Commentaires sur les comptes 2017

Pour la première fois de son histoire, l'Idiap termine l'exercice 2017 avec des revenus supérieurs à 11 millions.

Les subventions publiques, toujours inférieures à 50 % de l'ensemble des revenus, sont indispensables pour le bon fonctionnement de l'Institut et de son développement. L'augmentation de la part cantonale, soutenue par le Service des Hautes Ecoles de l'Etat du Valais, va permettre, en 2018, la création de deux nouveaux groupes de recherche qui devraient encore générer de nouvelles recettes.

La masse budgétaire liée aux projets, toujours croissante, provient notamment d'une nouvelle source de financement. Trois nouveaux projets sont maintenant financés par le Département de la Défense US, DARPA et IARPA.

Les frais de personnel (salaires + charges sociales) représentent environ trois quarts de la totalité des charges. Pour un franc cantonal, l'effet de levier actuel est supérieur à 5.5.

L'Idiap boucle ses comptes 2017 sur un bénéfice de 34'487 francs.

Subventions

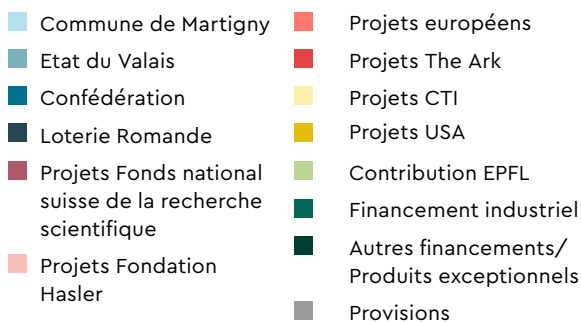
Confédération, Canton, Commune

(En milliers de francs suisses)

ANNÉES	2015	2016	2017	2018*
Confédération	2467	2424	2418	2406
Canton	1720	1720	2000	2000
Commune	700	700	700	700

*Budget

Répartition des sources de financement



Répartition des charges



Bilan (CHF)

ACTIF	31.12.2016	31.12.2017
Liquidités	5'313'117.83	5'306'132.52
Débiteurs	130'394.85	61'886.00
Actifs de régularisation et divers	434'238.84	769'938.24
TOTAL ACTIFS CIRCULANTS	5'877'751.52	6'137'956.76
Mobilier et matériel informatique	590'125.14	458'112.24
Participations financières	10'000.00	10'000.00
TOTAL ACTIFS IMMOBILISÉS	600'125.14	468'112.24
TOTAL ACTIF	6'477'876.66	6'606'069.00

PASSIF	31.12.2016	31.12.2017
Créanciers	213'106.26	472'526.46
Passifs de régularisation	3'747'835.83	2'517'150.17
Provisions	1'085'000.00	1'450'000.00
TOTAL FONDS ÉTRANGERS	5'045'942.09	4'439'676.63
Capital	40'000.00	40'000.00
Réserve Fonds de Recherche	0.00	700'000.00
Réserve spéciale	1'200'000.00	1'200'000.00
Résultat reporté	179'687.15	191'934.57
Résultat net	12'247.42	34'457.80
TOTAL FONDS PROPRES	1'431'934.57	2'166'392.37
TOTAL PASSIF	6'477'876.66	6'606'069.00

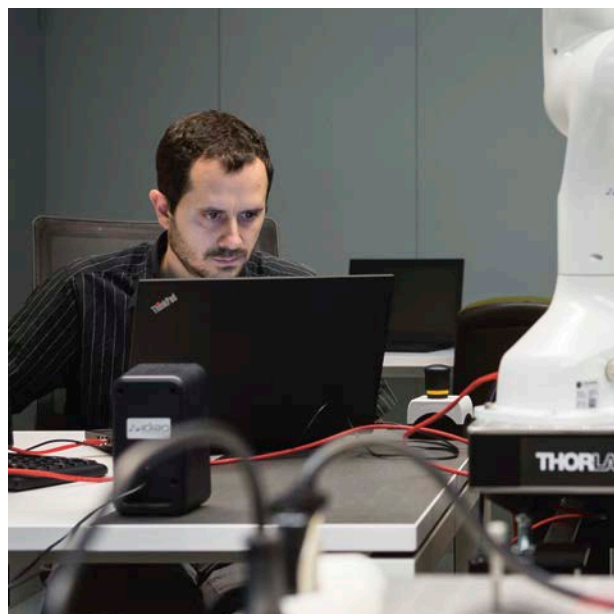


recherche

22 — 27

MEMMO – Une mémoire du mouvement pour qu'un robot puisse s'adapter rapidement

Pour chacun d'entre nous, enjamber un obstacle en travers de notre chemin est tout à fait naturel. Mais pour un exosquelette ou un robot quadrupède, cela relève plutôt de l'exploit. Pour maintenir son équilibre et le contrôle, il va devoir réaliser un rapide calcul de ses options. Une réponse qui tarde à arriver ou un mauvais choix sélectionné entraîne inexorablement sa chute. La solution : intégrer dans ses programmes des souvenirs d'expériences similaires de maintien d'équilibre. Et justement, l'Idiap se penche sur la question avec le programme européen MEMMO.



MEMMO (Memory of Motion, www.memmo-project.eu/) est un grand projet européen de 4 millions d'euros. D'une durée de 48 mois, il réunit un consortium d'intervenants académiques et industriels qui s'attèlera à trouver des solutions pour « donner » de la mémoire aux robots.

Sous la coordination d'une unité indépendante du CNRS (Centre national de la recherche scientifique), le Dr Sylvain Calinon, chercheur senior et responsable du groupe « Robot Learning & Interaction », va mener les équipes de l'Idiap. Les autres partenaires académiques sont les Universités d'Edimbourg et d'Oxford, ainsi que l'Institut Max-Planck de Tübingen.

Le consortium va au delà des labos universitaires, et intègre des partenaires industriels venus de toute l'Europe avec chacun leurs savoir-faire, mais aussi besoins spécifiques. PAL Robotics, de Barcelone, est spécialisé dans la création de robots humanoïdes ; Wandercraft, installé à Paris, développe des exosquelettes spécialement conçus pour les personnes paraplégiques. Airbus

s'est également joint au projet : le constructeur aéronautique requiert des robots anthropomorphiques pour réaliser des opérations dans les carlingues d'avion. Le britannique Costain Group souhaite quant à lui développer des robots quadrupèdes pour inspecter les canalisations.

Quels sont les défis du projet ? Il est difficile pour un robot d'apprendre quand il ne peut faire qu'un nombre limité d'expériences dans le monde réel. Cet apprentissage est d'autant plus crucial pour les robots à forme humanoïde qui seront certainement amenés à effectuer des tâches variées dans des environnements différents. Heureusement, une partie de ces expériences peut être réalisée en simulation, où le robot imagine plusieurs péripéties qui peuvent lui arriver dans un environnement virtuel, et essaie d'y faire face. Les expériences réussies sont ensuite analysées et conservées sous forme de mémoire de mouvement.

Grâce aux recherches menées dans le projet MEMMO, les robots précités seront capables de mieux gérer les changements d'environnement et les imprévus.

Martigny au cœur d'un réseau européen d'étude sur le langage

TAPAS (Training Network on Automatic Processing of Pathological Speech, www.tapas-etn-eu.org) n'est pas un projet comme les autres. Soutenu par un dispositif de financement très sélectif mis en place par l'Union Européenne, ce programme de recherche sur la parole chapeauté par l'Idiap mise sur la mobilité internationale et intersectorielle des chercheurs venus de tous les horizons. Un défi de taille et ambitieux qui propulse l'institut valaisan au cœur de la science de pointe « made in Europe ».

L'idée du projet est née de la collaboration de longue date entre le D^r Mathew Magimai-Doss, chercheur senior à l'Idiap et coordinateur de TAPAS, et la D^r Heidi Christensen de l'Université de Sheffield. Pour les accompagner dans cette aventure d'envergure et répondre aux différents objectifs de recherche, douze institutions académiques et industrielles, ainsi que neuf organisations partenaires ont intégré le consortium, dont l'Idiap est le chef d'orchestre.

Un projet de taille exceptionnelle et un financement exceptionnel : TAPAS a intégré le réputé programme de subventionnement « Marie Skłodowska-Curie Actions » (MSCA). Ce dernier encourage particulièrement les réseaux de formation dans des domaines scientifiques à la pointe de l'innovation misant sur la mobilité tant géographique qu'intersectorielle. Ce dont n'est pas peu fier le D^r Magimai-Doss : « Les financements de type MSCA sont très recherchés et impliquent une forte concurrence entre les projets soumis. »

Parole altérée : la nécessité d'une approche multidimensionnelle

De nombreuses pathologies peuvent affecter la capacité des hommes à s'exprimer. Parmi les plus fréquentes, on compte les maladies d'Alzheimer et de Parkinson, ou encore les accidents vasculaires cérébraux. Les personnes concernées vivent souvent une forme de mise à l'écart, voire d'exclusion.

Le projet TAPAS propose d'étudier la problématique sous trois angles principaux : la détection d'une altération du langage, les outils d'évaluation et de thérapie et la réhabilitation. Le D^r Magimai-Doss et ses partenaires ont imaginé une approche multidisciplinaire qui va de la recherche la plus fondamentale jusqu'aux

applications les plus concrètes. Le but commun de toutes ces approches, selon le coordinateur de TAPAS, « comprendre comment une pathologie affecte l'émission et la réception du langage ».

Quinze doctorants, quinze facettes du traitement de la parole

Dans le cadre du projet, un consortium de cliniciens, de groupes de recherche et de partenaires industriels va engager quinze doctorants répartis sur différents sites européens. En plus de leur cursus scientifique de trois ans, les doctorants recevront des formations théoriques et pratiques, allant de l'éthique et des questions légales relatives aux données à une formation en entrepreneuriat. L'objectif scientifique, lui, sera d'étudier les différentes facettes de l'altération de la parole et son incidence sur la communication.

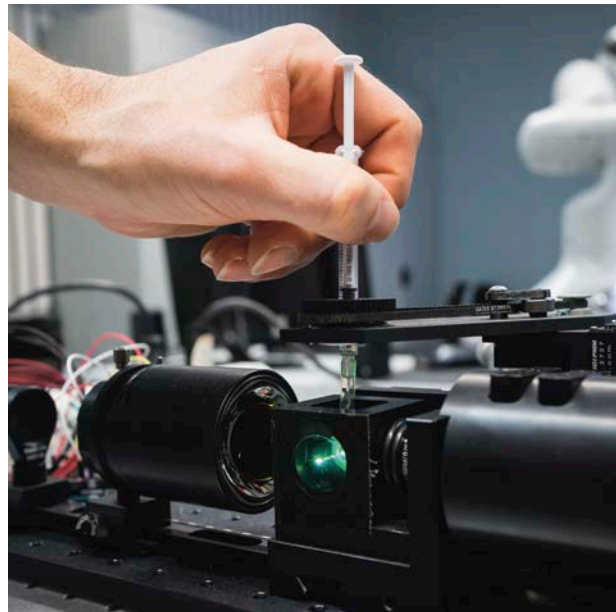
Un premier groupe d'étudiants se focalisera sur la mesure de l'intelligibilité des personnes souffrant d'une pathologie. Comme les méthodes traditionnelles d'évaluation de la parole humaine ne donnent pas de résultats satisfaisants en cas d'handicap, il s'agit par exemple de développer des nouveaux scores d'intelligibilité.

Un autre groupe de doctorants utilisera le langage et particulièrement les prémices des problèmes d'élocution afin de détecter des maladies dégénératives et d'en mesurer l'évolution.

Finalement, la question de comment aider au mieux les personnes concernées par une pathologie est au cœur d'une autre thématique de recherche. Ainsi, la modélisation du langage des enfants va permettre de créer un outil thérapeutique sous forme de jeu.

L'Idiap est dans le move

Une nouvelle plateforme inédite pour capter ce qui bouge et valider les résultats de recherche



Pourquoi cette nouvelle plateforme MMD (Multi-Modal-Data) ?

« Il y a un vide dans l'acquisition des données reproductibles provenant d'une source en mouvement, de capteurs en mouvement ou avec un éclairage qui change. Cette idée de plateforme est née lors d'une discussion entre chercheurs. Le besoin se faisait sentir dans différents groupes de l'Idiap », explique le Dr Michael Liebling qui est à la tête du groupe « Computational Bioimaging ». C'est en collaboration avec deux autres chercheurs seniors de l'Idiap, Sylvain Calinon et Sébastien Marcel, que Michael Liebling dépose une demande de subside au Fonds national suisse de la recherche scientifique (FNS) fin 2015. Celle-ci est acceptée au courant de l'année 2016.

La plateforme MMD vient s'ajouter à un dispositif déjà existant, comme la très performante plateforme BEAT qui traite des données biométriques.

Quel est le but de ce type de plateformes ?

« Lorsqu'un chercheur utilise des algorithmes dans ses recherches, comment peut-il être sûr que les résultats sont justes ? Est-ce que ces résultats vont être reproductibles à l'autre bout du monde ? », se demande Michael Liebling. Car, en effet, l'une des conditions nécessaires pour valider un résultat de recherche est que celui-ci puisse être reproduit.

L'Idiap soutient ces efforts de validation par la qualité de ses bases de données annotées qui permettent de reproduire des expériences avec des algorithmes établis et d'offrir la possibilité d'évaluer et comparer la qualité de nouveaux algorithmes.

Construction d'un système de microscopie made in Valais

La plateforme MMD imaginée et réalisée à l'Idiap comprend « un système de microscopie de type open-SPIM dédié à l'acquisition d'images en 3D, pour lequel il a fallu construire des pièces sur mesure ». Les polymécaniciens de la base aérienne de Sion ont non seulement réalisé les pièces mais leurs suggestions ont aussi permis d'éviter quelques pièges lors de l'assemblage.

Grâce à ce nouveau système de microscopie qui offre des possibilités de créer des points de comparaison, les chercheurs en bio-imagerie computationnelle peuvent valider leurs algorithmes. Ceux-ci reconstituent, par exemple, les images d'un cœur battant avec une précision inégalée. Michael Liebling, très enthousiaste, se réjouit de toutes les nouvelles possibilités qu'offre cette plateforme pour ses recherches et celles de ses collègues.

Pas loin de lui, Sylvain Calinon et son équipe ont placé un dispositif de capture d'images en 3D au bout de la main de l'un des deux bras robotiques qu'offre la plateforme. Le but est de reconstruire le mouvement du bras à partir des images captées depuis le bout de la main. Est-ce que les algorithmes utilisés seront efficaces ? MMD permettra de répondre à ces questions et bien d'autres.

Construire des bases de données dans des domaines qui en manquent, pour ensuite les partager, est un autre objectif visé. Avec cette nouvelle plateforme issue d'un environnement créatif et de plusieurs collaborations, les chercheurs de l'Idiap sont encore une fois à la pointe du mouvement.

L'Intelligence Artificielle sous le feu des projecteurs de la recherche

L'apprentissage automatique fait salle comble

L'apprentissage automatique, soit le « champ d'études qui donne aux ordinateurs la capacité d'apprendre sans être explicitement programmés » n'est pas quelque chose de nouveau : l'Américain Arthur Samuel avait déjà esquissé une première définition en 1962. Plus d'un demi-siècle plus tard, l'Intelligence Artificielle (IA) est (re)devenue un sujet d'intérêt pour la recherche, et au-delà, une question centrale pour la société. François Fleuret, responsable du « Machine Learning group » à l'Idiap et enseignant à l'EPFL spécialisé dans les réseaux neuronaux, organise chaque année depuis 2012 le « Swiss Machine Learning Day ». Une journée qui a pour but de rassembler les chercheurs suisses autour de la thématique de l'apprentissage automatique, et qui ne cesse de conquérir année après année un public toujours plus grand. En novembre dernier, victime de son succès, une des plus grandes salles du SwissTech Convention Center à Lausanne a été nécessaire pour accueillir les nombreux invités.

Torch, le couteau suisse de l'apprentissage automatique développé à l'Idiap

Afin de résoudre des questions complexes impliquant de nombreuses données, les algorithmes d'apprentissage automatique requièrent un support très flexible. En 2001, l'Idiap répond à ce besoin en développant Torch : un maillage de bibliothèques de « deep learning », idéal pour simuler des réseaux neuronaux. Près de 20 ans plus tard, Torch est devenu un outil 100% opérationnel, utilisé notamment par Facebook. Ce dernier a d'ailleurs développé un descendant direct de Torch, PyTorch, adopté entre autres par Uber et de nombreux autres chercheurs à travers le monde.



L'Idiap et la HES-SO Valais/Wallis main dans la main

Ouvert au public et à la communauté scientifique, les chercheurs des deux instituts valaisans organisent des workshops en commun sur l'Intelligence Artificielle. Afin de stimuler la coopération et les échanges, les deux instituts planifient deux sessions par année. Ces initiatives, gratuites et disponibles en ligne, se veulent être des plateformes universelles de dissémination de savoir. La société Klewel, une spin-off de l'Idiap, les met d'ailleurs à disposition du public, avec la possibilité de les visionner orateur par orateur.

La thématique du premier workshop fut dédiée aux questions liées à la reproductibilité des recherches scientifiques. Initiée par le Dr Sébastien Marcel, responsable du groupe de recherche « Biometrics Security and Privacy », cette question brûlante d'actualité a été abondamment discutée et débattue lors de cette journée. Un invité de marque, Pr Pierre Vanderghyest, vice-président pour l'éducation, est venu présenter la position de l'EPFL et ses solutions pour répondre à ce défi.



Le deuxième workshop fut organisé par le professeur David Ginsbourger, responsable du groupe « Uncertainty Quantification and Optimal Design » à l'Idiap. Cette fois-ci le focus s'est fait sur les Small Data et la problématique des recherches comptant peu de données, à l'image de celles dans le domaine des maladies rares, des catastrophes écologiques, climatiques ou financières. Même les chercheurs en robotique de l'Idiap sont confrontés aux obstacles des Small Data. Un sujet original, et qui touche à des modèles de mathématiques pures.

300 personnes à Martigny sur le thème de « L'IA dans tous ses états ».

Pour la 17^e édition du Forum Economique Rhodanien, les représentants des secteurs privés et publics des régions traversées par le Rhône se sont réunis à Martigny à l'hôtel Vatel le 15 septembre 2017. Son but : favoriser les échanges d'idées et soutenir des opportunités de réalisations entre des acteurs économiques de Suisse et de France voisine. La journée, sponsorisée par Business Valais et la Ville de Martigny, avait pour thème central le traitement des innovations technologiques induites

par l'IA. Le directeur de l'Idiap, le professeur Hervé Boulard, a partagé son point de vue sur la question lors de sa présentation « l'Intelligence Artificielle, entre réalité et fiction, entre recherche et propagande ». Les chercheurs seniors de l'Idiap Sébastien Marcel et Jean-Marc Odobez ont également pris part aux tables rondes dédiées à la cybersécurité et à la robotique. Cette rencontre a permis de beaux échanges entre la science, les instances politiques, économiques et le grand public.

réseau

28 — 33

La 6^e édition de l'International Create Challenge

Ce programme en immersion complète de trois semaines (www.createchallenge.org) a permis à des jeunes entrepreneurs de faire éclore leurs idées en des prototypes commerciaux. Organisé du 30 août au 20 septembre 2017 à l'Institut de Recherche Idiap de Martigny, il a réuni des participants venus de partout (Brésil, Ukraine, Chine, Argentine). Le jury de l'ICC, composé d'une dizaine de personnes de renommée internationale provenant du monde de la recherche, du capital-risque et de l'innovation, a décerné les prix aux équipes suivantes qui remportent chacune CHF 5000.- en cash :

- **Forensic Voice Recognition System** (Srikanth Madikeri, Idiap) développe un outil facilitant l'identification de personnes basée sur la biométrie vocale. Le projet gagne par ailleurs l'accès gratuit à l'incubateur The Ark.
- **Robobay** (Yaroslav Pishta, Ukraine) est une plateforme web permettant à ses clients de déployer leurs propres robots de trading en s'interfaçant au marché de la crypto-monnaie.
- **TOP DOWN** (Renato Rodrigues, Brésil) développe une technologie de monitoring, de fréquentation et de paiement qui facilite l'utilisation des transports publics.

« Ce challenge est possible grâce à la fidélité des experts qui accompagnent les participants. »

L'ICC, outil de transfert technologique qui a fait ses preuves

L'International Create Challenge (ICC) est à l'origine de sociétés telles que Recapp IT, Biowatch, EyeWare ou Anemomind. Pour la volée 2017, des négociations sont en cours avec plusieurs équipes pour intégrer l'incubateur The Ark sur le site d'IdeArk à Martigny.

François Foglia, directeur adjoint de l'Idiap et créateur de l'ICC, tient à souligner le rôle des experts qui participent depuis six ans à cette aventure. Pendant chaque jour du challenge, l'un d'eux vient partager gracieusement ses connaissances ou coacher les équipes d'entrepreneurs en herbe.

Christophe Saam (P&TS) est présent chaque année pour expliquer les règles de la propriété intellectuelle. Gaspard Couchepin, avocat au barreau, notaire et docteur en droit, vient présenter les différents types de sociétés. Laurent Bischof et Guillaume Dubret (Polytech Ventures) viennent exposer le fonctionnement des ventures capitalistes. Nadine Reichenthal, une autre experte et chargée de cours en stratégie à l'Université de Lausanne, vient parler de business canvas et de value proposition. Jordi Montserrat (Venture Kick) partage son expérience dans le démarrage des start-up. De leur côté, Jung Park (IMD), Joël Rossier, Daniel Rühle (TheArk), Jasmine Abarca-Golay (InnoPark), Patrick Biro (Venture Kick) et Stéphane Marchand-Maillet (Université de Genève) font aussi partie de ces fidèles qui apportent chaque année leur soutien aux participants ou contribuent aux travaux du jury.

Se tourner vers le grand public, une vocation pour l'Idiap

Carnotzet scientifique : la recherche en toute convivialité

Gratuit et ouvert à tous, le Carnotzet Scientifique Valaisan propose des réunions qui mettent le public et les experts invités sur un pied d'égalité. Un panel d'experts s'est retrouvé à la librairie La Liseuse à Sion le 10 mai 2017 pour débattre autour du thème « Serious Game : utiliser une approche ludique pour résoudre des problèmes très sérieux ». Le Carnotzet Scientifique était l'occasion idéale pour échanger avec des citoyens autour d'un verre, en toute décontraction. Le professeur Daniel Gatica-Perez, directeur du « Social Computing Group » de l'Idiap, a d'ailleurs présenté le smartphone comme un outil souvent sous-estimé de contribution citoyenne. Un constat issu de ses recherches à l'Idiap, au croisement de l'informatique, des médias sociaux, de l'apprentissage automatique et des sciences sociales.

Pepper, une star valaisanne sur YouTube

Pepper, un des robots de l'Idiap, a été invité à s'exprimer au TEDx Martigny le 8 septembre 2017. Devant une salle comble, il présente, non sans humour, la raison de sa présence : « Les organisateurs n'ont pas trouvé un être humain suffisamment idiot pour parler d'intelligence artificielle. Et, comme pour de nombreuses choses que le bipède éclairé n'aime pas faire, c'est moi qui m'y colle. » Une prestation des plus remarquées et qui continue de faire des émules sur Internet : son intervention mise en ligne sur Youtube comptabilise déjà plus de 2400 vues.

Début novembre, l'attachant robot a été l'atout charme de la cérémonie de remise des prix culturels de l'Etat du Valais. Pepper a en effet volé la vedette en interviewant à distance la comédienne valaisanne Mali Van Valenberg, qui ne pouvait être présente ce jour là. Son écran de contrôle a permis à l'artiste d'adresser directement ses remerciements par écran interposé. Très vite adopté par les autres lauréats présents, Pepper a même eu le droit de poser pour la photo de groupe ! Une réussite pour les chercheurs, et une jolie découverte pour le public.

Venez visiter l'Idiap !

L'opportunité de visiter l'Idiap est offerte à tous chaque 2^e mercredi du mois de 17 à 18 heures (ou sur demande). François Foglia, directeur adjoint de l'institut, y présente les activités des groupes de recherche et répond volontiers aux questions des visiteurs les plus curieux.

Activités pour le jeune public

Le futur de la recherche se conjugue à tous les genres

Depuis 2012, l'institut accueille pendant une matinée des enfants dans le cadre de « futur en tout genre ». En novembre dernier, dix-neuf enfants âgés de 11 à 12 ans sont venus découvrir les différentes facettes du métier de chercheur, de développeur, gestionnaire de réseaux ou de projets. En marge des ateliers organisés à leur attention, une rencontre est organisée avec Nao et Pepper, les robots de l'Idiap. Cette initiative du Secrétariat valaisan à l'égalité et à la famille offre la possibilité aux enfants d'envisager leur avenir sans barrière de genre. En ouvrant le champ des possibles, ils peuvent s'autoriser plus aisément à être en contact avec leur propre élan intérieur. Il faudra cependant attendre quelques années pour savoir si l'un d'entre eux ou elles aura eu un coup de cœur pour les sciences.

Hérisson sous gazon à Charrat

Ce festival open air dédié aux enfants de 3 à 14 ans a lieu une fois par an à Charrat, et pas moins de 7000 personnes viennent fouler le gazon le temps d'un week-end de juin. Présent chaque année, l'Idiap ne rate jamais cette occasion de présenter un atelier ludique où petits et grands peuvent découvrir et interagir avec des robots. Une jolie vitrine pour présenter les recherches de l'institut qui s'est spécialisé dans la reconnaissance des signaux, qu'ils soient auditifs, visuels, sociaux ou comportementaux. Le jeune public est particulièrement ravi d'échanger avec Nao, le petit robot de l'Idiap.



Un rayonnement sans frontière pour les spin-off de l'Idiap

B iowatch (www.biowatch.ch), Recapp (www.recapp.ch), Eyeware (eyeware.tech) ou Koemei. Quatre start-up, quatre success stories, un point commun : les labos de l'Idiap. Coup de projecteur sur une année 2017 synonyme de réussite pour ces spin-off et l'institut.

Toutes les portes s'ouvrent à Biowatch

Tout commence en 2014. Matthias Vanoni, alors en thèse de doctorat, se lance dans un pari fou : créer un module qui ouvre toutes les portes, même virtuelles, grâce à un capteur placé sur les veines du poignet (une technique développée dans les laboratoires de Martigny). Après une première phase de soutien via le programme d'International Create Challenge (www.createchallenge.org), le projet Biowatch rencontre un véritable succès en 2017 : il est le lauréat du premier prix du Swiss Fintech Convention à Genève, et remporte le prix Visa's Everywhere Initiative (qui a pour but d'identifier le « next big thing » en matière de paiement), et ce en marge du réputé World Mobile Congress de Barcelone. La spin-off a décroché plusieurs autres prestigieuses distinctions, notamment de Thales, de LVMH, d'Accenture ou encore du Crédit Suisse. Autre nouvelle qui ravit les potentiels investisseurs : Biowatch a reçu l'année dernière le label CTI, attestant de sa solidité et maturité malgré son jeune âge.

Une vague valaisanne au SWISS ICT Award 2017

À Lucerne en novembre dernier, deux spin-off de l'Idiap se sont particulièrement distinguées lors de la soirée des Swiss ICT Awards, un événement récompensant chaque année les produits informatiques et des services fondés sur les TIC les plus innovants en Suisse. Recapp, spécialisée dans l'enregistrement et la

retranscription audio dernière génération, s'est vue décerner le « newcomer award » et s'est hissée à la 2^e place de leur top 100. La spin-off n'est pas une inconnue sur le marché suisse : elle collabore entre autres avec le Parlement valaisan (projet MediaParl) ou encore Swisscom. Enfin Eyeware, une solution inventive d'interaction visuelle entre les personnes handicapées et les ordinateurs (prix de l'ICC'2015), a réussi à être nominée lors de ces Swiss ICT Awards.

Un rachat qui va donner de la voix

La société suisse Crealogix, spécialiste en solution FinTech pour les banques, s'est offerte la technologie développée par Koemei. Créée en 2010, cette spin-off de l'Idiap s'est construite à partir des travaux de recherche sur la reconnaissance vocale, l'analyse de texte et du langage naturel pour classer automatiquement des contenus. Cette solution était notamment utilisée par plusieurs universités dans le cadre des formations supérieures.

biowatch

 **eyeware**

[rec]app

Les alumni : un trait d'union riche entre la recherche et l'industrie

Passer des labos à l'industrie, il n'y a qu'un seul pas. Et plusieurs collaborateurs de l'Idiap n'ont pas hésité à le franchir, avec succès. À titre d'exemples, Milos Cernak est aujourd'hui directeur technique chez Logitech, ou encore Alexandros Lazaridis, devenu directeur du Speech Processing dans le département « Artificial Intelligence and Machine Learning » chez Swisscom.

L'écosystème dynamique de l'Idiap offre un terrain fertile aux collaborations entre les chercheurs. Au-delà d'une simple ambiance, c'est un véritable état d'esprit entrepreneurial qui anime les couloirs. Et qui va même au-delà, comme pour Hugues Salamin et Yann Rodriguez, tous deux chez KeyLemon.

Coopération et proximité, deux facteurs clés de réussite à l'Idiap

Alors doctorant entre 2007 et 2009 en « Social Signal Processing », Hugues rencontre l'un des fondateurs de KeyLemon, Yann Rodriguez. À cette époque, ce dernier avait terminé son doctorat et travaillait déjà avec des développeurs sur son projet. Le courant passe parfaitement entre les deux chercheurs, et nul ne soupçonne que ce premier contact scellera leurs avenir, mais aussi celui de la spin-off KeyLemon.

Hugues part pour quelque temps en Ecosse, mais il revient à l'institut en 2014 en tant que développeur pour reprendre ensuite les commandes du bureau de transfert technologique. Un poste qui lui permet d'acquies une solide expérience de développement de projets avec les industries, et d'enrichir son réseau professionnel. Un profil qui ne laissera pas indifférent KeyLemon, alors en pleine expansion, et qui viendra le chercher en 2017.

Devenu depuis responsable produit, le Valaisan loue les avantages de cette proximité créée et nourrie à l'Idiap : « Tout le monde se connaît, et cela facilite grandement les collaborations car les attentes de part et d'autre sont claires. Nous avons besoin d'une évaluation extérieure de notre solution pour téléphone portable, et c'est un postdoc de l'Idiap qui s'en est chargé. Cela nous a permis d'avancer plus vite. »

Autres points majeurs appréciables selon Hugues : il a été possible de trouver une place de travail en Valais dans son domaine, faisant ainsi de l'institut un vivier attrayant de recrutement pour les sociétés du canton.



visages

34 — 39

Du MIT à Martigny, l'odyssée d'un chercheur globe-trotteur passionné de langage

Le chercheur américain James Henderson a intégré avec brio l'équipe de l'Idiap en automne dernier. Menant le groupe de recherche « Natural Language Understanding », il se concentre entre autres sur le traitement sémantique de textes et discours pour en améliorer la traduction automatique. Gros plan sur une carrière aussi brillante que nomade dédiée à la science du langage.

Pour James Henderson, tout a commencé il y a près de trente ans au MIT (Massachusetts Institute of Technology). Depuis, le goût de la recherche internationale ne l'a plus quitté. Il est passé entre autres dans les universités d'Exeter, de Genève ou encore d'Edinburgh. Il a également travaillé au sein des laboratoires de recherche de Xerox à Grenoble.

Dans son domaine, le Dr Henderson est un pionnier du genre. Au tout début des années 2000, il mit au point avec succès des techniques novatrices mêlant l'analyse des réseaux de neurones et la compréhension du langage naturel, un des plus grands défis de l'intelligence artificielle.

James Henderson reprend aujourd'hui le flambeau du professeur Andrei Popescu-Belis, qui a rejoint l'HEIG-VD à Yverdon. Le nouveau groupe de recherche s'inscrit dans la continuité des travaux menés dans le traitement du langage naturel à l'Idiap, et ces derniers porteront précisément sur la traduction, l'indexation, la classification de textes et leurs implications.

Révéler les opinions à partir de Facebook

Une implication lexicale ou textuelle établit une relation entre deux mots ou deux textes. Par exemple, « ceci est un chat » implique que le chat est un animal, une évidence même. Mais pour faire de tels liens, un ordinateur doit d'abord passer par un processus d'apprentissage. Passionné par ces questions, le Dr Henderson souhaite utiliser l'étude des implications du langage pour synthétiser des opinions. À l'aide d'un modèle extrayant le sens des phrases à l'aide de ces implications, il sera possible d'établir un consensus à partir



d'une grande variété d'opinions divergentes. Le chercheur américain y voit une application directe dans le domaine des sondages d'opinion : au lieu de passer par des questionnaires ou des enquêtes téléphoniques, il sera possible de déterminer l'opinion d'un groupe de personnes en extirpant et analysant leurs écrits sur les réseaux sociaux.

Au-delà des labos, un état d'esprit créatif

Pour James Henderson, l'Idiap offre à la fois un environnement intellectuellement stimulant et une grande liberté pour les chercheurs. « Les séminaires hebdomadaires, où nous présentons nos propres recherches, sont des plus intéressants. Tout comme les échanges entre les doctorants et postdoctorants des différents groupes ». Un état d'esprit construit sur le partage d'idées qui génère une émulation positive selon lui, créant dès lors une communauté créative, dynamique et propice aux découvertes.

Mélanie Huck et Guillaume Clivaz fêtent leur premier anniversaire au sein de l'Idiap

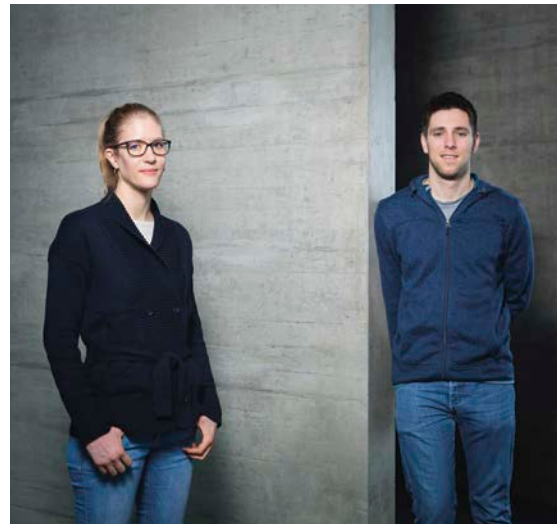
Il y a maintenant plus d'un an que Mélanie et Guillaume ont démarré leur nouvelle aventure au sein du groupe d'ingénieurs de recherche et de développement de l'Idiap. Ces deux jeunes aux parcours très différents viennent rejoindre ce groupe qui se distingue par sa grande diversité.

De l'architecture aux constructions virtuelles

Après avoir accompli des études d'architecture à l'EPFL et travaillé un certain temps dans l'architecture, Mélanie veut se lancer un nouveau défi : développer des logiciels dans son domaine de prédilection. Pour cela elle rejoint la Haute Ecole d'Ingénierie d'Yverdon pour y effectuer un nouveau diplôme. Elle s'aperçoit très rapidement que les possibilités sont beaucoup plus vastes que ce qu'elle avait imaginé et décide de changer d'horizon. Ces pas l'amènent à l'Idiap qu'elle intègre en février 2017. Cette Veveysanne, qui pratique aussi le taekwondo au niveau international, apprécie la nouveauté et les challenges. Depuis son arrivée, elle a travaillé sur quatre projets aussi différents que passionnants. Mélanie aime le traitement d'image et la 3D et elle imagine l'élaboration d'un grand logiciel comme une construction. Décidément l'architecture n'est jamais très loin.

Un environnement dynamique en Valais

Après des études en microtechnique et un master en robotique à l'EPFL, Guillaume, de son côté, enchaîne deux expériences temporaires dans l'industrie des procédés chimiques ainsi que dans le domaine des matériaux de carbone pour l'industrie de l'aluminium. Ses recherches d'emploi le mènent ensuite à l'Idiap et au groupe de recherche et de développement. Depuis ses débuts à l'institut martignérain, il contribue à un vaste projet qui lui a permis de nombreux échanges et collaborations. Mais l'un des buts de ce tout jeune développeur est de se plonger dans le machine learning et en particulier de se frotter aux techniques du deep learning. En dehors de son travail, Guillaume prend aussi le temps de s'investir dans la fanfare de sa région



en jouant du saxophone et fait également partie du club de tennis local.

Le groupe de recherche et développement de l'Idiap puise sa force dans la diversité des intérêts et les parcours éclectiques de ses ingénieurs. Ils peuvent ainsi répondre avec originalité et efficacité à la grande variété des demandes émanant du monde industriel et des chercheurs au sein même de l'Idiap. Ils participent ainsi pleinement à l'effort de transfert technologique de la recherche à l'industrie, l'une des trois missions de l'Idiap. Il est à noter, avec fierté, que de cette même équipe sont issus de nombreux ingénieurs ayant rejoint l'industrie suisse ou internationale, consolidant les liens qui unissent l'Idiap avec le monde économique. Et ce n'est pas là l'un de ses moindres succès.

Souhaitons donc à Mélanie et Guillaume un parcours aussi enrichissant qu'épanouissant.

Une juriste au carrefour de l'Intelligence Artificielle

Arrivée en mars 2017, Marie-Constance Kaiflin est un vrai concentré d'énergie. Première juriste engagée à l'Idiap, la jeune femme se sent comme un poisson dans l'eau dans ce milieu de chercheurs et de développeurs. Un trait d'union pétillant, moderne et nécessaire entre la science et le droit.

Avant de travailler à Martigny, cette Genevoise détentrice d'un Master en droit s'était déjà aventurée professionnellement en Valais.

Heureuse de ses premières expériences à Sion, d'abord dans une petite étude d'avocat et ensuite dans différents départements à l'Etat du Valais, elle accepte la proposition originale qui lui est faite à l'Idiap. « Créer son poste est une belle opportunité », souligne Marie-Constance Kaiflin. Propulsée dans le domaine de l'Intelligence Artificielle, elle se dit ravie de travailler dans une ambiance à la fois « dynamique, entrepreneuriale et cosmopolite ».

Le couteau suisse des négociations de contrat

Ses tâches sont aussi variées que nombreuses : elle revoit les contrats de travail et gère les règlements internes, elle analyse et rédige les contrats de service avec les partenaires industriels et les conventions de recherches dont les projets sont financés par Innosuisse (l'Agence suisse pour l'encouragement de l'innovation). Pour le grand projet européen TAPAS, dont l'Idiap est coordinateur (voir p. 24), elle négocie l'accord de consortium avec les onze autres parties, dont des instituts académiques ou de recherche ainsi que des partenaires industriels.

Petit à petit, les demandes de conseils juridiques affluent de partout, y compris de start-up de la plateforme technologique IdeArk situées dans le même bâtiment que l'Idiap. Elle répond avec le même entrain aux questions sur des contrats, des contraintes



légales soulevées par la présence d'un robot dans un supermarché finlandais ou encore des recherches juridiques sur l'éventuelle protection d'un dictionnaire de traduction français-suisse-allemand par le droit d'auteur.

Juriste : se former non-stop au quotidien

Comment a-t-elle fait pour s'adapter aussi rapidement ? « J'ai des collègues formidables ! J'ai beaucoup appris grâce à eux », s'exclame-t-elle. La première personne qu'elle cite est Florent Monay, responsable du transfert technologique, avec qui elle partage le même bureau.

L'apprentissage se fait par la pratique mais aussi en ligne. Afin de se former à la nouvelle législation européenne sur la protection des données, elle suit le module de droit du programme « CAS Data protection, Privacy and Biometrics » proposé par UniDistance. Avec la montée en puissance de la digitalisation et la révolution du Big Data, de nouvelles questions de droit et d'éthique se posent. Elles requièrent d'être étudiées de façon plus pointue. C'est pourquoi Marie-Constance participe aussi à la création d'un comité d'éthique. Quel est au fond le secret de sa réussite ? « Je m'intéresse à ce qui m'entoure, j'aime le mouvement », explique-t-elle.

Prix Idiap

Interne

Chaque année, l'Idiap décerne deux prix destinés à ses doctorants. Le premier récompense une recherche, le second une publication. Pour l'attribution du prix Idiap de la Recherche, le candidat est évalué par une commission interne sur la base de cinq critères: ses publications, son esprit d'équipe, son engagement, son sens de la communication et son autonomie. Pour le prix de la Publication, une première sélection est effectuée par les chercheurs permanents de l'institut parmi les travaux dont l'auteur principal est un doctorant Idiap. Les membres du Comité d'accompagnement notent ensuite, séparément et de façon anonyme, les écrits choisis.

En 2017, le prix de la Recherche a été attribué à **James Newling** pour son autonomie et sa motivation exceptionnelle dans ses recherches dans le domaine du clustering (k-means et k-medoids). Celui de la Publication a été remis à **Gülcan Can** pour son excellent article scientifique intitulé « Codical Maya Glyph Segmentation: A Crowdsourcing Approach ».



Gülcan Can, groupe de Social Computing



James Newling, groupe de Machine Learning

Externe

Cette année, l'Idiap tient à relever la très brillante participation de ses chercheurs lors des conférences internationales. La qualité de leur recherche a été récompensée par plusieurs prix.

8th International Joint Conference on Natural Language Processing
 Novembre 27 – décembre 1, Tapei, Taiwan
 Best Paper Award

Nikolaos Pappas et Andrei Popescu-Belis

Multilingual Hierarchical Attention Networks for Document Classification

Intelligent Service Robotics (ISR) journal,
 2017 Best Paper Award

Sylvain Calinon

A Tutorial on Task-Parameterized Movement Learning and Retrieval

20th International Conference on Artificial Intelligence and Statistics
 Mai 2017, Fort Lauderdale, Floride, USA
 Best Paper Award

James Newling et François Fleuret

A Sub-Quadratic Exact Medoid Algorithm

Thèses achevées

Cinq étudiants ont achevé leur thèse (EPFL) en 2017 :
Pierre-Edouard Honnet, Pedro Henrique Oliveira Pinheiro,
Olivier Canévet, Marzieh Razavi et Gülcan Can.

Intonation Modelling for Speech Synthesis
and Emphasis Preservation

Pierre-Edouard HONNET

Janvier 2017

Directeur de thèse: P^r Hervé Bourlard

Membres du jury: D^r Jean-Marc Vesin, D^r Philip N. Garner, P^r Jean-Philippe Thiran, D^r Junichi Yamagishi et D^r Antonio Bonafonte

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

Marzieh RAZAVI

Août 2017

*Directeurs de thèse: P^r Hervé Bourlard
et D^r Mathew Magimai-Doss*

*Membres du jury: P^r S. Sússtrunk, D^r K. Knill, P^r M. Davel
et P^r J.-Ph. Thiran*

Large-Scale Image Segmentation with
Convolutional Networks

Pedro Henrique OLIVEIRA PINHEIRO

Janvier 2017

Directeurs de thèse: P^r Hervé Bourlard et D^r Ronan Collobert

*Membres du jury: P^r Jean-Philippe Thiran, P^r Rob Fergus,
D^r Cordelia Schmid et D^r Mathieu Salzmann*

Visual Analysis of Maya Glyphs via
Crowdsourcing and Deep Learning

Gülcan CAN

Septembre 2017

*Directeurs de thèse: P^r Daniel Gatica-Perez
et D^r Jean-Marc Odobez*

*Membres du jury: D^r Denis Gillet, P^r Alberto del Bimbo,
P^r Rolf Ingold et P^r Jean-Philippe Thiran*

Object Detection with Active Sample
Harvesting

Olivier CANÉVET

Février 2017

Directeur de thèse: D^r François Fleuret

*Membres du jury: P^r Pascal Frossard, P^r Gilles Blanchard,
P^r Raphael Salzmann et D^r Mathieu Salzmann*



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

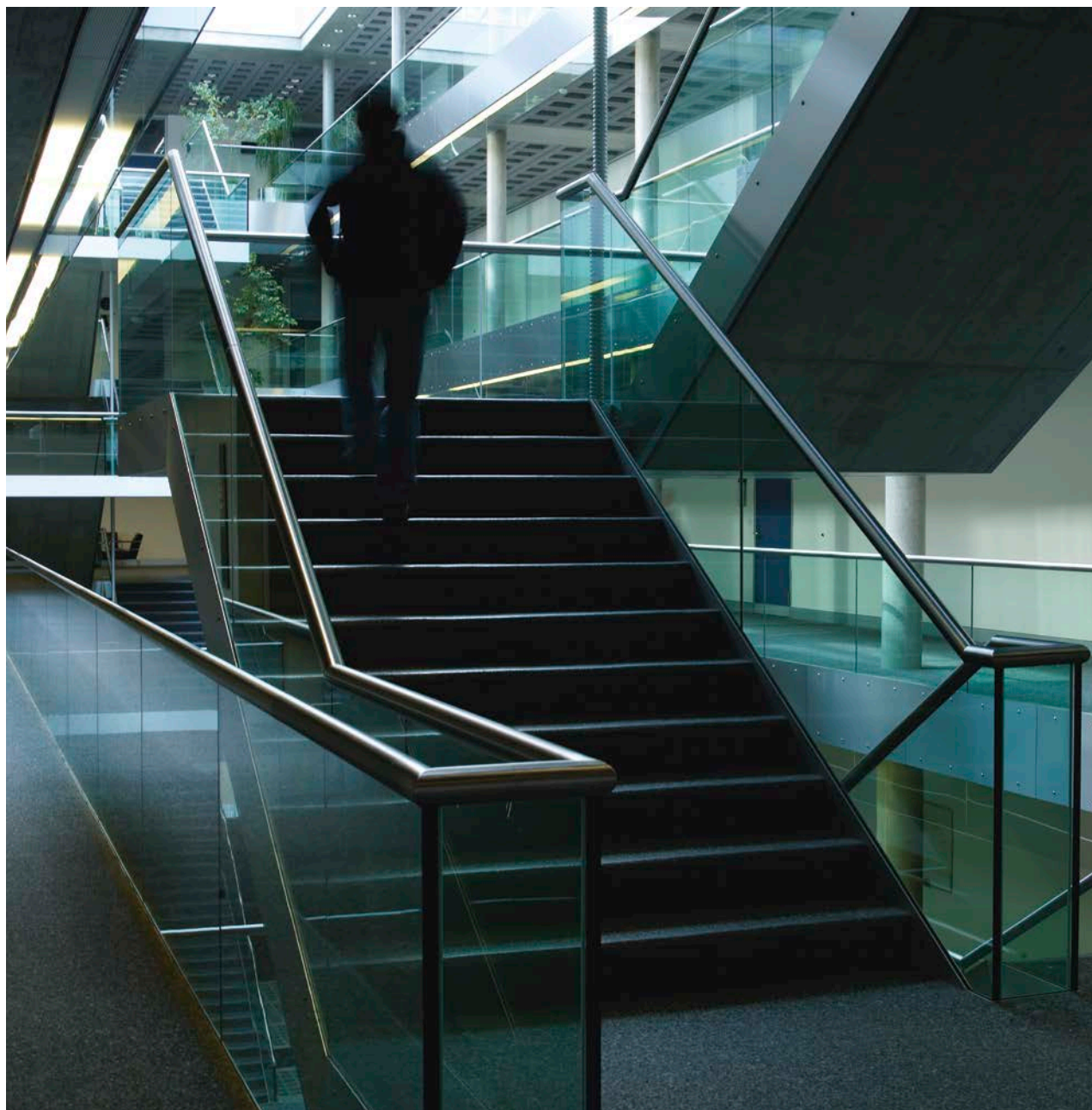
T +41 27 721 77 11

F +41 27 721 77 12

M info@idiap.ch

www.idiap.ch

40



scientific inserts

I – XL

Speech and Audio Processing

Overview

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

Group overview

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

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

Key scientific outputs

Idiap has always significantly contributed to both Hidden Markov Model (HMM) and Deep Neural Network (DNN) based approaches applied in acoustic modelling for various speech processing tasks. Use of techniques from HMM and HMM-DNN based speech recognition in HMM and HMM-DNN based speech synthesis resulted in a unified approach to speech recognition and synthesis. The group was well placed to

take full advantage of recent advances in new architectures of deep learning, studied in particular through *PyTorch* and other open source frameworks. Advances in Automatic Speech Recognition (ASR) are usually researched through *Kaldi* toolkit, now used by most of the international speech community.

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

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

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

Automatic speech recognition

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

→ Robust parametrisation and acoustic modelling

We are still investigating new features (e.g., posterior-based features) and new acoustic models (new forms of hidden Markov models, such as KL-HMM, or artificial neural networks) that are more robust to noise and acoustic environments, as well as to speaker variability (e.g., accented speech, or dialect).

→ Cross-lingual and multi-lingual speech recognition

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

project (Summarization and domain-Adaptive Retrieval of Information Across Languages) aims at developing cross-lingual retrieval and summarization techniques that will work for any language in the world, given minimal resources to work with. In those contexts, we focus on investigating and exploiting fast acoustic model adaptation techniques in cross-lingual and multi-lingual scenarios. The resulting speech recogniser relies on a hybrid approach, where an artificial neural network acoustic model is boot-strapped using well-resourced data and adapted to the target language.

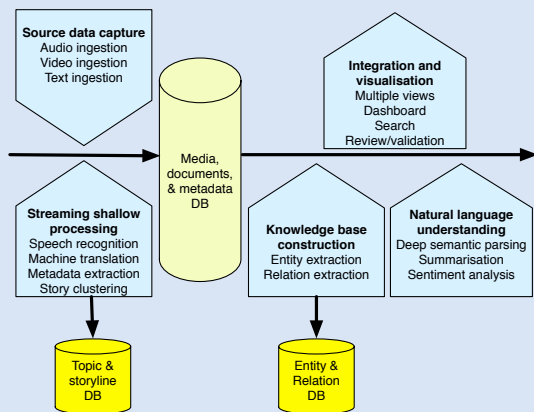


Figure 1
An overview of the EC H2020 SUMMA project, a current focus for multilingual speech recognition at Idiap.

→ **Swiss languages**

We continuously improve our speech recognisers for Swiss German and Swiss French and also apply the most recent advances in speech technology employing deep neural networks. The recognisers are also used in commercial products of the Idiap spinoff *recapp IT AG*. Work on speech recognition for Walliserdeutsch one of the most difficult to understand of the Swiss dialects, was first published in 2014; the associated database is also available for download. Since 2015, we collaborate with *recapp IT AG* on a wider range of Swiss dialects towards the first commercial product that performs Swiss

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

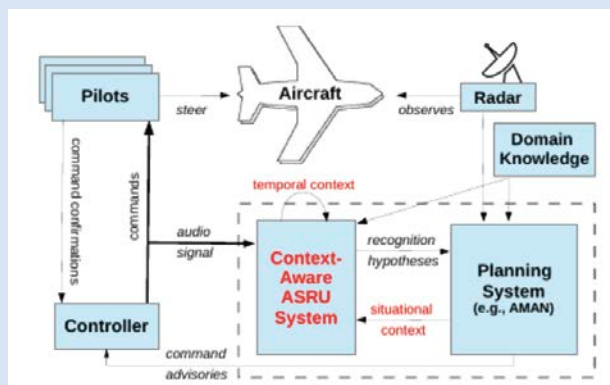
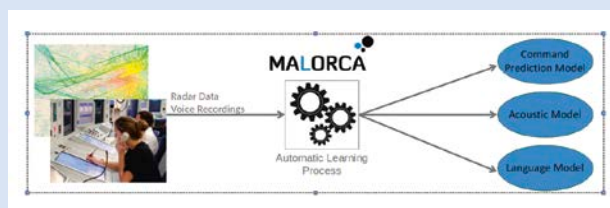


Figure 2
Up: Overview of the EC MALORCA project; semi-supervised adaptation of three models. **Down:** graphical representation of building blocks.

→ **Template-based ASR**

Idiap continues to work on template-based or exemplar-based ASR using posterior-based features in the context of projects like SNSF projects A-MUSE, PHASER, and PHASER-QUAD, and CTI project MultiVEO². In this context, there are currently two on-going efforts. The first focuses on development of "sparse" and "low-rank" HMM frameworks by establishing a link between statistical speech recognition formalism and the recent theories of sparse modelling and (probabilistic) principle component analysis (PCA). The framework integrates the advantages of template-based modelling within the HMM sequence matching scheme. The second effort focusses on development of an approach that seamlessly combines both statistical sequence model based ASR system and template-based ASR system into a single ASR system that can yield performance better than the best individual system.

→ **Lexicon development**

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

→ **Semi-supervised learning of acoustic and language model parameters**

This problem has been conducted at Idiap over several past years, as it aims to develop a technology to (1) automatically select speech data which is acoustically highly informative, or (2) to automatically assess recognition output with high confidence. This type of technology is essential to allow semisupervised learning of speech recognition models using unlabelled data. One of the specific architecture has been developed within the EC H2020 MALORCA project², combining both acoustic (i.e. speech) and situational (i.e. radar) information to build automatic speech recognition system for air-traffic controllers (i.e. controller-pilot communication). The MALORCA project, graphically represented in Figure 2 further focused on technologies to automatically adapt the developed system for new airports, while minimising human effort to manually transcribe speech data.

→ **Punctuation prediction**

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

→ **Failure analysis of ASR and HSR decoding channels**

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

Speech synthesis and coding→ **Text-to-speech synthesis (TTS):**

Although newer than ASR, TTS is now an established venture for the speech group at Idiap. TTS has been central to several projects, including: SIWIS³ (Spoken Interaction With Interpretation in Switzerland), D-BOX⁴ and SP2⁵ (SCOPES project on speech prosody). The TTS work at Idiap was largely aligned with the statistical synthesis trend, which uses the same technology as ASR. However, the group has tracked the recent developments in deep learning which will dominate future research. SIWIS aimed to do speech to speech translation in Swiss languages, including the concept of accents. Current work under MASS (Multilingual Affective Speech Synthesis) brings the concept of emotion into the speech synthesis, particularly via modelling of prosody.

→ **Speech coding**

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

² www.malorca-project.de

³ www.idiap.ch/project/siwis/

⁴ www.idiap.ch/scientific-research/projects/dbox

⁵ www.idiap.ch/scientific-research/projects/sp2

Speaker recognition and speech data analytics

In the context of the European SIIP project⁶, illustrated by Figure 3, the Speech and Audio Processing group has significantly improved their capabilities in suspect identification applicable to very large scale data. The SIIP technology has successfully passed three proof-of-concept and field-test events in 2016 and 2017, demonstrating the performance of the developed technology among the key stakeholders. The developed suspect identification solution can analyse not only lawfully intercepted calls, but also multiple types of social-media channels. SIIP has also developed a framework allowing to combine evidences extracted by different types of engines (i.e. inter-task engines such as language/accents/gender/age identification, or keyword-spotting) to eventually improve speaker identification.

As discussed in Section 8, and as part of the SNSF Project UniTS⁷, the group is also contributing to the Biometric Person Recognition group, including the development of countermeasures to detect attacks on speaker verification systems through forged speech samples.

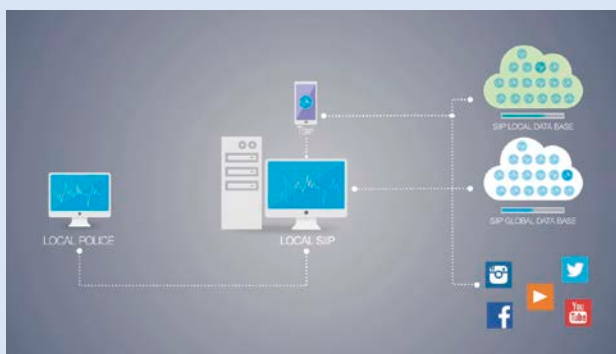


Figure 3

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

Forensic speaker verification and audio analysis

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

Large scale spoken query retrieval

- **Query-by-example spoken term detection (QbE-STD) based on subspace modelling and detection**
Mainly exploiting its knowhow in Deep Neural Networks (DNN) and sparse recovery modeling, Idiap continues its research efforts towards keyword spotting and spoken term detection with a focus on searching large audio archives using spoken queries. Over the past 12 months, Idiap has developed pioneering retrieval solutions relying on the characterization and detection of the low-dimensional subspace of the DNN phonetic posteriors.
- **Large scale spoken query indexing**
In 2017, Idiap further developed a powerful methodology for large-scale spoken query retrieval relying on hashing. Contribution of hashing is two-fold: (1) Compressing the information bearing data characteristic and (2) Identifying a data-driven symbolic space for effective search. Idiap hashing solution relies on DNN representation learning at sub-phonetic attribute level to facilitate cross-lingual applications.
- **Objective speech intelligibility assessment**
Over the past two years, Idiap has been actively developing novel methods to objectively assess intelligibility of speech signal. This work is now exploited in the context of HASLER Foundation project FLOSS² and the pathological speech processing activities discussed below.

Pathological speech processing

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

- **MoSpeedi (Motor Speech Disorder) SNSF Sinergia project:** aiming at characterizing phonetic speech planning and motor speech programming/execution and their impairments, in collaboration with Marina Laganaro, Faculty of Psychology and Educational Science, University of Geneva, Cécile Fougeron, Laboratoire de Phonétique et Phonologie, Paris 3-Sorbonne Nouvelle, and Frédéric Assal, Neurology, Geneva University Hospitals and Faculty of Medicine, University of Geneva.
- **EC TAPAS (Training Network on Automatic Processing of Pathological Speech)** coordinated by Idiap: targeting three key research problems, (1) detection (develop speech processing techniques for early detection of conditions that impact on speech production), (2) therapy (using newly-emerging speech processing techniques to produce automated speech therapy tools), and (3) assisted Living (redesigning current speech technology so that it works well for people with speech impairments and also helps in making informed clinical choices).

⁶ www.siip.eu

⁷ www.idiap.ch/scientific-research/projects/units

² www.idiap.ch/project/floss/

Sign language recognition and assessment

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

Sound localization and microphone array

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

Key publications

- [1] P.-E. Honnet, B. Gerazov, A. Gjoreski, and P. N. Garner, "Intonation modelling using a muscle model and perceptually weighted matching pursuit", *Speech Communication*, Vol. 97, 2018.
- [2] D. Ram, A. Asaei, and H. Bourlard, "Sparse Subspace Modeling for Query by Example Spoken Term Detection," to be published in *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, March 2018.
- [3] A. Asaei, C. Milos, and H. Bourlard, "Perceptual Information Loss due to Impaired Speech Production", *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, Issue: 99, August 2017.
- [4] M. Razavi, R. Rasipuram and M. Magimai-Doss, "Towards Weakly Supervised Acoustic Subword Unit Discovery and Lexicon Development Using Hidden Markov Models", *Speech Communication*, Volume 96, February 2018, Pages 168-183
- [5] S. Tong, P. Garner, and Bourlard, "An Investigation of Deep Neural Networks for Multilingual Speech Recognition Training and Adaptation", *Proc. of Interspeech*, Proceedings of Interspeech, Stockholm, Sweden, 2017.
- [6] S. Dey, P. Motlicek, S. Madikeri, and M. Ferras, "Exploiting Sequence Information for Text-Dependent Speaker Verification", *Proc. of 2017 IEEE Intl. Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, New Orleans, pages 5370-5374, 2017.

Machine Learning

Overview

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

Group overview

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

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

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

Key scientific outputs

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

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

Efficient Machine learning

→ Large-scale clustering

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

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

→ Sub-linear hard sample extraction

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

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

→ High-dimension similarity measures

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

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

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

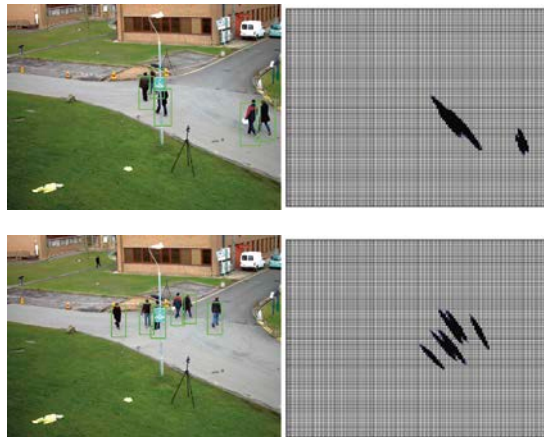


Figure 4
Person detection obtained with a deep neural network, and the corresponding occupancy map estimates.

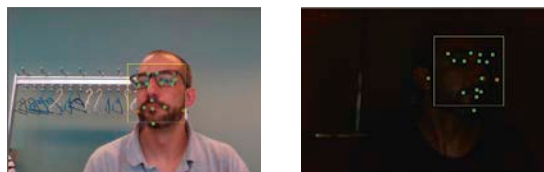


Figure 5
Face alignment examples with qualitative visibility estimation. The color from green to red indicates the visible confidence from high to low.

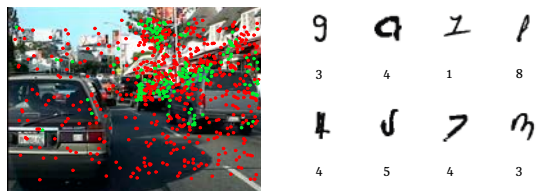


Figure 6
We have developed a novel strategy to find informative training examples in an image, that concentrates the computation on difficult images and parts of images to train an object detector (left), and on difficult and unusual examples to train an image classifier (right).

Deep learning

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

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

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

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

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

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

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

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

Key publications

- [1] J. Newling and F. Fleuret. K-Medoids For K-Means Seeding. In Proceedings of the international conference on Neural Information Processing Systems (NIPS), 2017.
- [2] S. Tulyakov, A. Ivanov, and F. Fleuret. Weakly Supervised Learning of Deep Metrics for Stereo Reconstruction. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), pages 1348–1357, 2017.
- [3] P. Baqué, F. Fleuret, and P. Fua. Deep Occlusion Reasoning for Multi-Camera Multi-Target Detection. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), pages 271–279, 2017.
- [4] T. Chavdarova and F. Fleuret. Deep Multi-Camera People Detection. In Proceedings of the IEEE International Conference on Machine Learning and Applications (ICMLA), pages 848–853, 2017.
- [5] J. Newling and F. Fleuret. A Sub-Quadratic Exact Medoid Algorithm. In Proceedings of the international conference on Artificial Intelligence and Statistics (AISTATS), pages 185–193, 2017. (Best paper award).

Social Computing

Overview

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

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

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

→ Key scientific outputs

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

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

Social media analytics and mobile crowdsourcing for cities and health

Our work in this domain spans several research lines. First, as part of the SNSF Youth@Night and Dusk2Dawn projects⁹ (Multidisciplinary Study of Young People's Going Out and Drinking Behaviors, in collaboration with Addiction Switzerland and the University of Zurich), we investigated the use of mobile crowdsourcing and social media analytics to characterize urban phenomena, both in Switzerland and other world regions.¹⁰ This included the automatic recognition of place ambiance from social media images using deep learning [1] (Figure 7), and automatic recognition of drinking activity from smartphone sensor data, for a population of 200 young volunteers in Switzerland.

Regarding mobile crowdsourcing and health, in the context of the Bites-n-Bits project¹¹ (Understanding Eating Routines in Context, in collaboration with Nestle Research Center), we developed a smartphone sensing app to collect everyday

life data about eating and drinking for a population of 120 college students, and used it to build an approach for recognition of eating occasions [2]. Furthermore, using visual social media (Instagram) as data source, we analyzed eating and drinking patterns in the context of the Social Food and People project¹². (supported by EPFL Integrative Food Science and Nutrition Center.)

Regarding mobile crowdsourcing for social innovation, we continued our work using the Civique platform¹³ (Mobile Data for Local Causes), which included its use for teaching students about humanitarian technologies¹⁴, and to collect multimedia data from Swiss heritage speakers living in the USA¹⁵ (in a collaboration with the University of Lausanne, supported by the UNIL-EPFL Program on Collaborative Research on Science and Society). Finally, our work in Latin America on mobile crowdsourcing for youth engagement on local urban issues was expanded through a collaboration with institutions in Mexico and Colombia [3]. Other collaborations using the Civique platform are under development.

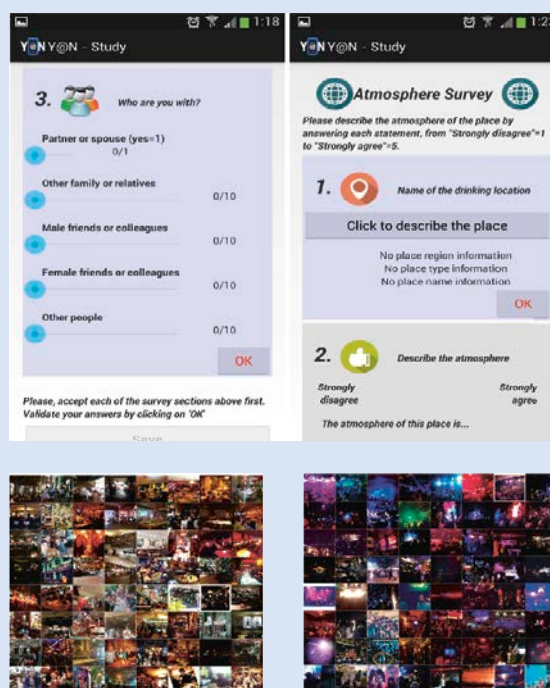


Figure 7

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

⁹ www.youth-night.ch

¹⁰ www.youtube.com/watch?v=71ht15VAoLw

¹¹ www.bitesnbits.org

¹² www.idiap.ch/en/allnews/what-does-instagram-reveal-about-food-consumption-in-switzerland

¹³ www.civique.org

¹⁴ www.idiap.ch/en/allnews/what-does-instagram-reveal-about-food-consumption-in-switzerland

¹⁵ www.idiap.ch/en/allnews/understanding-how-technology-can-revolutionize-humanitarian-work

¹⁵ wp.unil.ch/sina/

Ubiquitous interaction analytics

In the context of the SNSF UBImpressed project¹⁶ (Ubiquitous First Impressions and Ubiquitous Awareness), we have developed methods to analyze dyadic interactions in the workplace using multiple sensors (cameras, Kinect devices, microphone arrays, smartphones, and wearables), and to infer variables like perceived hirability and performance (Figure 8). This is joint work with the University of Lausanne and Cornell University, in partnership with Vatel Switzerland Hotel Management School. Specifically, we found connections between automatically measured nonverbal behavior (including speaking activity, prosody, head activity, and body motion) and impressions in job interviews and hospitality service encounters [4]. Some of this work will be transferred to industry in the context of a new CTI project.



Figure 8

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

Visual analysis of Maya hieroglyphs

In the context of the SNSF MAAYA project¹⁷ (Multimedia Analysis and Access for Documentation and Decipherment of Maya Epigraphy), we developed methods for visual analysis of hieroglyphs in ancient codices¹⁸. Specifically, we designed a methodology that included crowdsourced glyph localization and segmentation by non-experts for data labeling, and deep learning for visual recognition [5] (Figure 9). This included visualization of the output of deep networks for interpretation of the learned visual representations. Our work was covered by Horizons Magazine, the scientific magazine of the SNSF and the Swiss Academies of Arts and Sciences¹⁹.



Figure 9

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

Key publications

- [1] Y. Benkhedda, D. Santani, and D. Gatica-Perez, "Venues in Social Media: Examining Ambiance Perception Through Scene Semantics," in Proc. ACM Int. Conf. on Multimedia (MM), Mountain View, Oct. 2017.
- [2] J.-I. Biel, N. Martin, D. Labbe, and D. Gatica-Perez, "Bites'n'Bits: Eating Behavior in Context from Mobile Data," PACM on Interactive, Mobile, Wearable, and Ubiquitous Technologies (IMWUT), Vol. 1, No. 4, Dec. 2017.
- [3] S. Ruiz-Correa, D. Santani, B. Ramirez Salazar, I. Ruiz Correa, F. Alba Rendon-Huerta, C. Olmos Carrillo, B. C. Sandoval Mexicano, A. H. Arcos Garcia, R. Hasimoto Beltran, and D. Gatica-Perez, "SenseCityVity: Mobile Sensing, Urban Awareness, and Collective Action in Mexico," IEEE Pervasive Computing, Special Issue on Smart Cities, pp. 44–53, Apr.-Jun. 2017.
- [4] S. Muralidhar, M. Schmid Mast, and D. Gatica-Perez, "How May I Help You? Behavior and Impressions in Hospitality Service Encounters," in Proc. ACM Int. Conf. on Multimodal Interaction (ICMI), Glasgow, Nov. 2017.
- [5] G. Can, J.-M. Odobez, and D. Gatica-Perez, "Codical Maya Glyph Segmentation: A Crowdsourcing Approach," IEEE Trans. on Multimedia, published online Sep. 2017, Vol. 20, No. 3, pp. pp. 711–725, Mar. 2018.

¹⁶ www.idiap.ch/project/ubimpressed

¹⁷ www.idiap.ch/project/maaya

¹⁸ lab.idiap.ch/maaya/

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

Perception and Activity Understanding

Overview

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

Group overview

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

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

Key scientific outputs

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

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

Deep learning

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

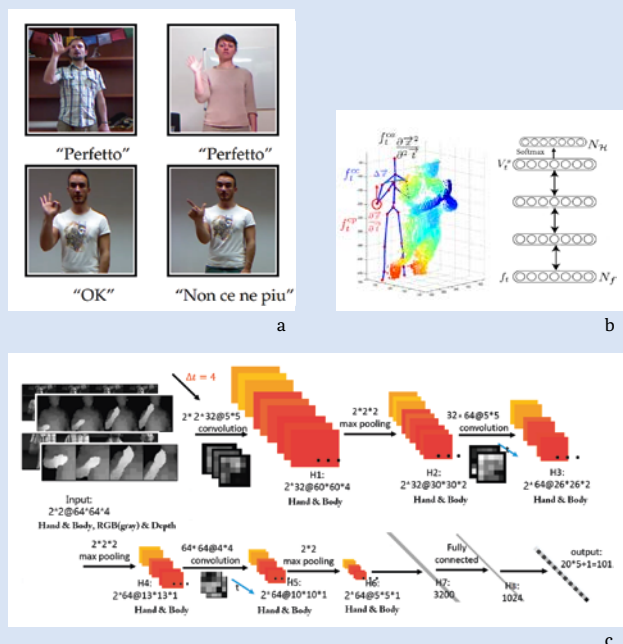


Figure 10

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

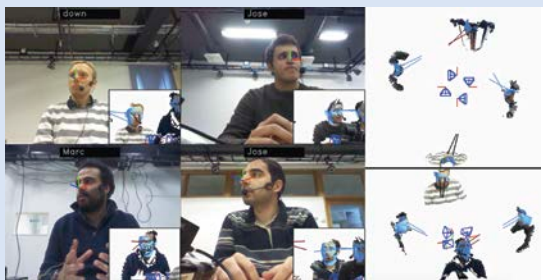
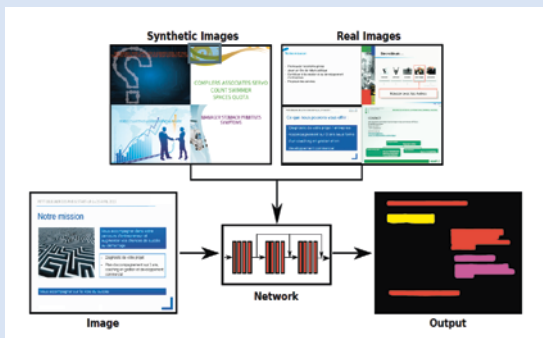


Figure 11

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

XII

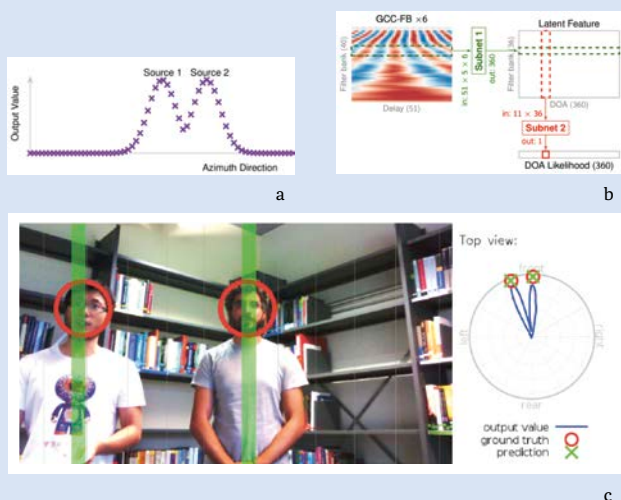


Figure 12

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

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

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

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

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

Multiple object tracking

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

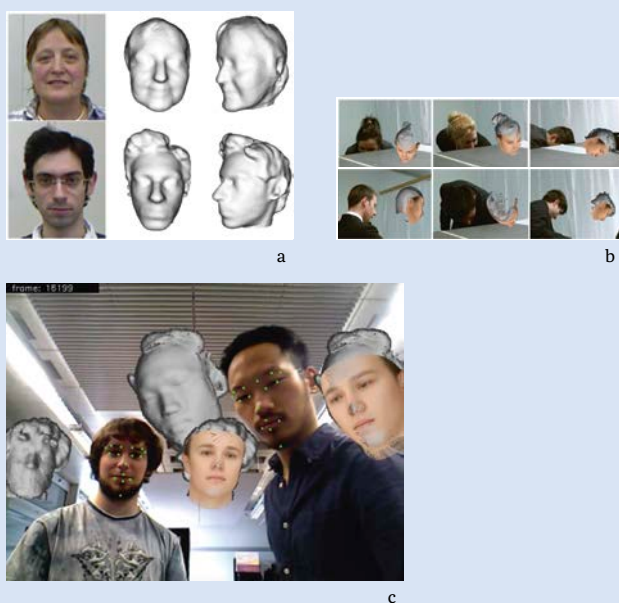


Figure 13
 Robust head pose tracking. **a)** Example of automatically re-constructed heads. **b)** Tracking results in adverse conditions (UBIImpressed registration desk set-up). **c)** Tracking results in an HRI situation (image taken from the Pepper robot).

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

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

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

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

Multimodal face and person diarization and naming

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

Key publications

- [1] Y. Yu, K. Funes and J.-M Odobez. HeadFusion: 360oHead Pose tracking combining 3D Morphable Mode and 3D Reconstruction IEEE Trans. Pattern Anal. Mach. Intell., accepted for publication, 2017.
- [2] G. Can, J.-M. Odobez, and D. Gatica-Perez. Codical Maya Glyph Segmentation: A Crowdsourcing Approach, IEEE Trans. on Multimedia, published online Sep. 2017, Vol. 20, No. 3, pp. pp. 711-725, Mars 2018.
- [3] N. Le and J.-M. Odobez. A Domain Adaptation Approach to Improve Speaker Turn Embedding Using Face Representation, In Int. Conf. on Multimodal Interactions (ICMI), 2017.
- [4] K. Funes and J.-M Odobez. Gaze estimation in the 3d space using rgb-d sensors: Towards head-pose and user invariance. Int. Journal of Computer Vision, 118(2):194-216, june 2016.
- [5] D. Wu, L. Pigou, P.-J. Kindermans, N. Le, L. Shao, J. Dambre, and J.-M. Odobez. Deep dynamic neural networks for multimodal gesture segmentation and recognition, IEEE Trans. Pattern Anal. Mach. Intell., 38(8):1583-1597, 2016.

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-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.

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

XIV

Key scientific outputs

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

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

Probabilistic models of movements and skills

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

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

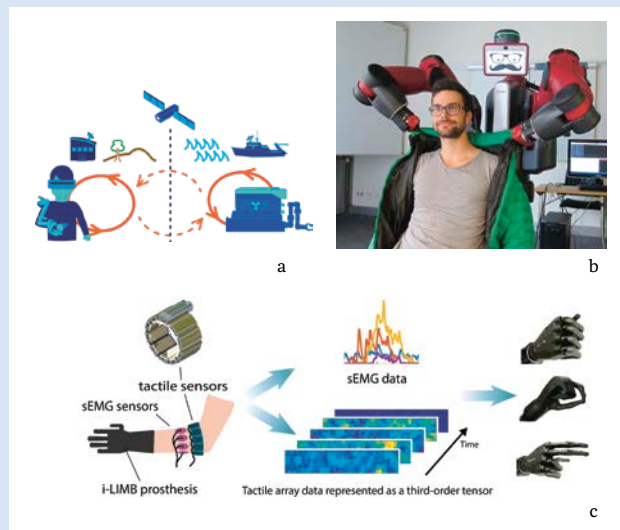


Figure 14

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

A distinctive view of optimal control

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

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

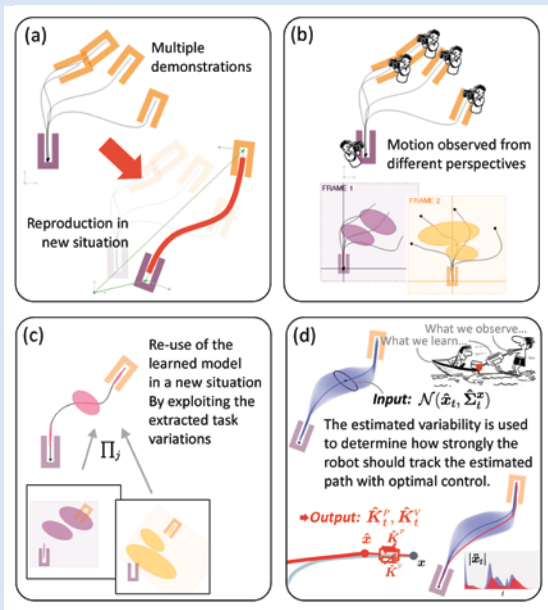


Figure 15

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

Geometry-aware statistical learning and control

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

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

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

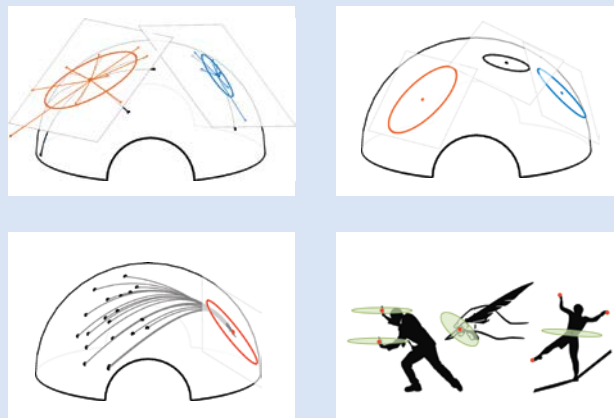


Figure 16

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

XV

Key publications

- [1] Pignat, E. and Calinon, S. (2017). Learning adaptive dressing assistance from human demonstration. *Robotics and Autonomous Systems*, 93, 61-75.
- [2] Jaquier, N., Connan, M., Castellini, C. and Calinon, S. (2017). Combining Electromyography and Tactile Myography to Improve Hand and Wrist Activity Detection in Prostheses. *Technologies*, 5:4, Special Issue on Assistive Robotics.
- [3] Zeestraten, M.J.A., Havoutis, I., Silvério, J., Calinon, S. and Caldwell, D.G. (2017). An Approach for Imitation Learning on Riemannian Manifolds. *IEEE Robotics and Automation Letters (RA-L)*, 2:3, 1240-1247.
- [4] Calinon, S. (2016). A Tutorial on Task-Parameterized Movement Learning and Retrieval. *Intelligent Service Robotics*, 9:1, 1-29.
- [5] Berio, D., Calinon, S. and Leymarie, F.F. (2017). Generating Calligraphic Trajectories with Model Predictive Control. In *Proc. of the 43rd Conf. on Graphics Interface*, pp. 132-139.

Uncertainty Quantification and Optimal Design

Overview

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

Group overview

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

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

Key scientific outputs

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

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

Bayesian optimization and emulation with Gaussian Process models

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

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

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

Computer experiments for the quantification and the reduction of uncertainties

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

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

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

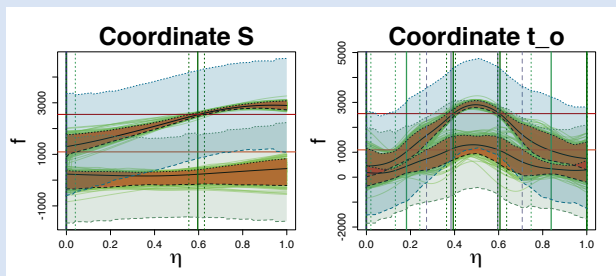


Figure 17
Estimating profile optima on a coastal flooding BRGM application

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

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

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

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

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

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

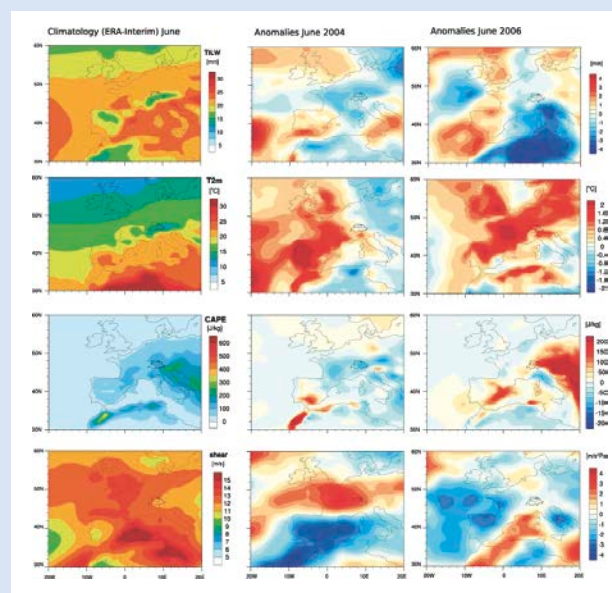


Figure 18
Environmental variables and their anomalies, used for hail prediction

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

Computational Bioimaging

Overview

Key publications

- [1] A Poisson regression approach to model monthly hail occurrence in Northern Switzerland using large-scale environmental variables, Erica Madonna, David Ginsbourger and Olivia Martius, in: Atmospheric Research, 203:261–274, 2018.
- [2] Estimating orthant probabilities of high dimensional Gaussian vectors with an application to set estimation, Dario Azzimonti and David Ginsbourger, in: Journal of Computational and Graphical Statistics, Accepted author version posted online: 03 Aug 2017.
- [3] Non-parametric warping via local scale estimation for non-stationary Gaussian process modelling, Sébastien Marmin, Jean Baccou, Jacques Liandrat and David Ginsbourger, in: Wavelets and Sparsity XVII, International Society for Optics and Photonics, 2017.
- [4] On uncertainty quantification in hydrogeology and hydrogeophysics, Niklas Linde, David Ginsbourger, James Irving, Fabio Nobile and Arnaud Doucet, in: Advances in Water Resources, 110:166–181, 2017.

XVIII

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

Group overview

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

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

Key scientific outputs

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

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

Direct inversion algorithm for focal plane scanning optical projection tomography

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

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

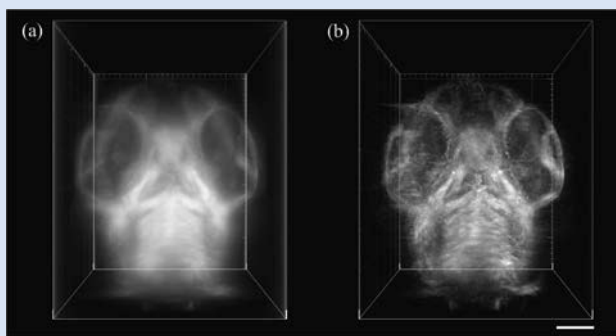


Figure 19

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

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

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

- Two 7-axis backdrivable robotic arms (Panda, Franka Emika), computer-driven
- Pneumatic gripper system
- Two-color Open-SPIM light-sheet microscope (20-water immersion microscope object, 488 et 561 nm laser excitation for green and red fluorescence imaging), computer-driven
- Dedicated data processing, storage and serving backend (4GPU NVidia K80 (19968 GPU cores), 8CPU Intel (112 CPU cores), 40TB storage)

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

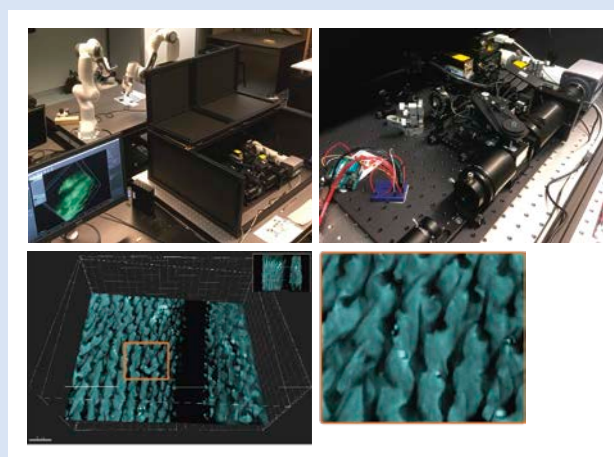


Figure 20

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

Key publications

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

Biometrics Security and Privacy

Overview

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

Group overview

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

- Biometric recognition: We investigate and develop new biometrics based recognition algorithms, notably for face, voice, and vein biometric modalities.
- Presentation attack detection: We look for new and better ways of detecting presentation attacks on face, voice, and vein biometric recognition systems.
- Biometric template protection: We research effective methods of preserving both the security of biometric recognition systems and the privacy of their users by protecting the biometric models ("templates") that are employed by the system for recognition purposes.

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

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

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

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

Key scientific outputs

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

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

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

Face and speaker recognition

We leveraged prior work on distribution modelling for part-based face recognition using session variability modelling techniques. Session variability modelling aims to explicitly model and suppress detrimental within-class (inter-session) variation. We examined two techniques to do this, inter-session variability modelling (ISV) and joint factor analysis (JFA), which were initially developed for speaker recognition. Next, we explored Total Variability modelling (TV), so called i-Vectors originally proposed for speaker recognition, for the task of face recognition. We also developed recently a scalable formulation of Probabilistic Linear Discriminant Analysis (PLDA). PLDA is a probabilistic model that has been shown to provide state-of-the-art performance for both face and speaker recognition.



Figure 21

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

Heterogeneous face recognition

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

Presentation attack detection

One important aspect of biometric systems is their reliability not only when assaulted by impostors, but also under different types of attacks. One possible security treat is presentation attacks (aka spoofing attacks): an action of outwitting a biometric sensor by presenting a counterfeit biometric evidence of a valid user (Figure 22). It is a direct attack to the sensory input of the biometric system and the attacker does not need previous knowledge about the recognition algorithm. Most of the biometric modalities are not resistant to presentation attacks: a biometric system is usually designed to only recognise identities without concern whether the sample comes from a live person or not. Despite the existence of very sophisticated biometric systems nowadays, the task of implementing presentation attack detection (PAD aka anti-spoofing) schemes for them has attracted much less attention.

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

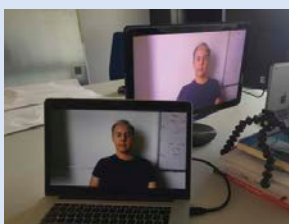


Figure 22
Illustration of video and audio presentation attacks.

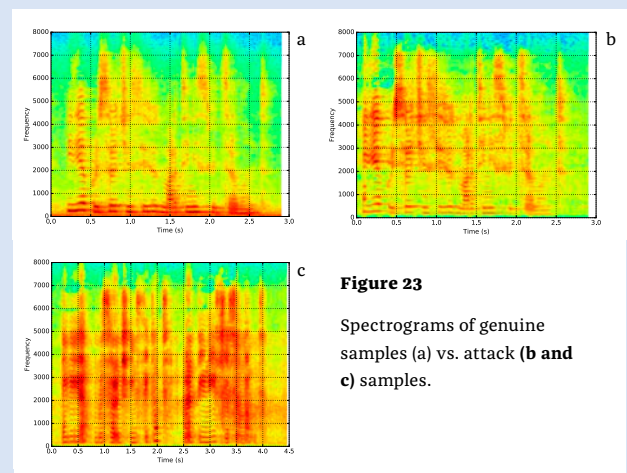


Figure 23
Spectrograms of genuine samples (a) vs. attack (b and c) samples.

Remote photoplethysmography

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

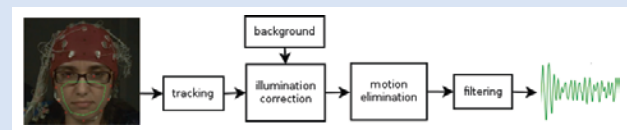


Figure 24
Illustration of remote photoplethysmography: colors from the video signal are filtered to produce an estimation of the heart beat signal.

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

Vein biometrics

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

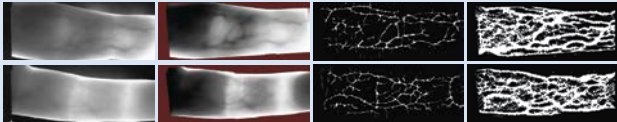


Figure 25

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

Swiss Centre for Biometrics Research and Testing

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

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

- Improving the security of the BEAT platform: an external IT security company conducted an audit to identify vulnerabilities of the infrastructure, allowing our engineers to strengthen the platform.
- Enriching the BEAT platform: our researchers and engineers developed Deep Learning and speech processing technologies to enrich the BEAT platform with new content. More particularly our engineers developed a new feature to execute GPU simulations at the heart of modern AI tools.
- Testing: our researchers conducted the independent testing of biometrics products from SMEs (e.g., KeyLemon now acquired by AMS) and the evaluation of a prototype for a large corporation.

Key publications

- [1] A. Mohammadi, S. Bhattacharjee and S. Marcel, "Deeply Vulnerable – a study of the robustness of face recognition to presentation attacks", *IET Biometrics*, 1-13, 2017.
- [2] P. Korshunov and S. Marcel, "Impact of score fusion on voice biometrics and presentation attack detection in cross-database evaluations", *IEEE Journal of Selected Topics in Signal Processing*, 695-705, 2017.
- [3] H. Muckenhirn, P. Korshunov, M. Magimai-Doss and S. Marcel, "Long-Term Spectral Statistics for Voice Presentation Attack Detection", *IEEE/ACM Transactions on Audio, Speech and Language Processing*, 2017.
- [4] S. Bhattacharjee and S. Marcel, "What you can't see can help you: extended-range imaging for 3D-mask presentation attack detection", *IEEE International Conference of the Biometrics Special Interest Group (BIOSIG)*, 2017.
- [5] A. Anjos, L. El-Shafey and S. Marcel, "BEAT: An Open-Science Web Platform", *International Conference on Machine Learning*, 2017.

Natural Language Processing

Overview

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

Group overview

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

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

Key scientific outputs

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

Document-level machine translation

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

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

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

XXIII

Example 1
Source: nach einföhrung dieser politik [...] die politik auf dem gebiet der informationstechnik [...]
Reference: once the policy is implemented [...] the information technology policy [...]
Baseline MT: after introduction of policy [...] the politics in the area of information technology [...]
Consistent MT: after introduction of policy [...] the policy in the area of information technology [...]
Example 2
Source: 欺詐性旅行或身份證件系指有下列情形之一之任何旅行或身份證件
Reference: Fraudulent travel or identity document; shall mean any travel or identity document
MT: 欺詐性 travel or identity papers. 系指 have under one condition; any travel, or identity document

Figure 26

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

²⁰ Supported by the COMTIS SNSF Sinergia project, 2010–2013, see www.idiap.ch/project/comtis.

²¹ Supported by the MODERN SNSF Sinergia project, 2013–2017, see www.idiap.ch/project/modern.

Text analysis for multilingual document representations

We develop methods for text and multimedia information retrieval, classification and recommendation, based on the analysis of content and sentiment, and on the networked structure of (social) media repositories. This work was supported by the SUMMA EU project²².

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

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

XXIV

Key publications

- [1] Pappas N. & Popescu-Belis A. (2017) – Multilingual Hierarchical Attention Networks for Document Classification. Proceedings of the 8th International Joint Conference on Natural Language Translation (WMT), Copenhagen, Denmark. Processing (IJCNLP), Taipei, Taiwan. Best paper award.
- [2] Pu X., Pappas N. & Popescu-Belis A. (2017) – Sense-Aware Statistical Machine Translation using Adaptive Context-Dependent Clustering. Proceedings of the 2nd Conference on Machine Translation (WMT), Copenhagen, Denmark.
- [3] Pu X., Mascarell L. & Popescu-Belis A. (2017) – Consistent Translation of Repeated Nouns using Syntactic and Semantic Cues. Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL), Valencia, Spain.
- [4] MiculicichWerlen L. & Popescu-Belis A. (2017) – Using Coreference Links to Improve Spanish-to-English Machine Translation. Proceedings of the EACL Workshop on Coreference Resolution Beyond OntoNotes (CORBON), Valencia, Spain.
- [5] Pappas N. & Popescu-Belis A. (2017) – Explicit Document Modeling through Weighted Multiple-Instance Learning. Journal of Artificial Intelligence Research (JAIR), vol. 58, p. 591–626.

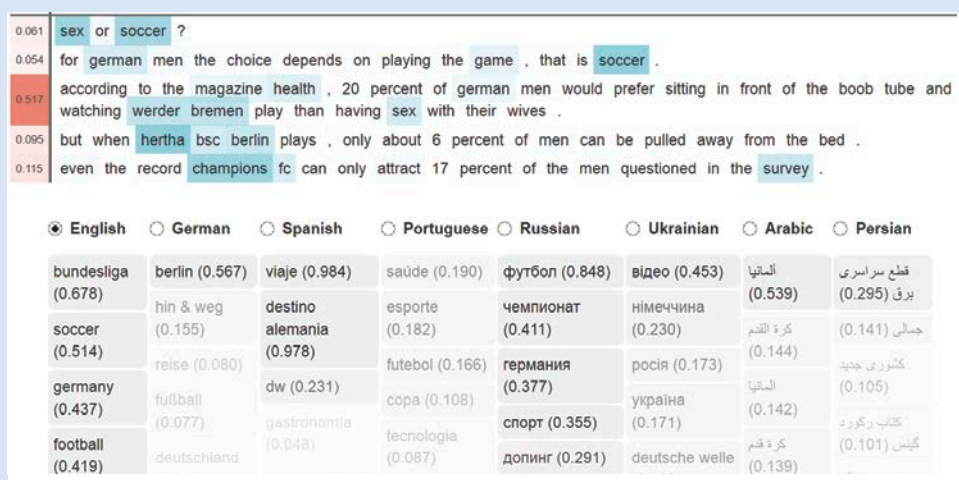


Figure 27
Multilingual topic labeling in 8 languages of a news article from Deutsche Welle.

²² See www.summa-project.eu.

Natural Language Understanding

Overview

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

Group overview

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

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

Key scientific outputs

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

Neural Network Architectures for NLP Tasks

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

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

xxv

Vector-Space Models of Entailment

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

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

Key publications

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

Active and Granted Projects in 2017

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

Projects in Progress during 2017

Projects Funded by European and International Agencies

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

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

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

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

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

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

ACTIVE AND GRANTED PROJECTS IN 2017

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, Acquirer AG, Bitplane AG, Leica Microsystems Cms Gmbh, 4D-Nature ImAGING Consulting, S. L., Centre Europeen de Recherche En Biologie et Medecine

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

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

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

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

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

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

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

Projects Funded by Swiss Agencies

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

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

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

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

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

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

ACTIVE AND GRANTED PROJECTS IN 2017

Name	MASH-2 (Massive Sets of Heuristics for Machine Learning II)
Funding	Hasler Foundation
Coordinator	Idiap Research Institute
Duration	2013.07.01 – 2017.09.30
Name	MEMUDE (Multi-view Detection with Metric-learning for Deep Network Fusion)
Funding	Hasler Foundation
Coordinator	Idiap Research Institute
Duration	2017.06.01 – 2019.05.31
Name	SETVISU (Learning and visualizing dangerous regions in multivariate parameter spaces)
Funding	Hasler Foundation
Coordinator	Idiap Research Institute
Duration	2016.12.01 – 2017.08.31
Name	GPUS (Acquisition d'un cluster de calcul dédié à l'Intelligence Artificielle)
Funding	Loterie Romande
Coordinator	Idiap Research Institute
Duration	2017.09.01 – 2018.08.31
Name	VALAIS+ (Une plateforme pour mieux connaître l'espace de vie du canton)
Funding	Loterie Romande
Coordinator	Idiap Research Institute
Duration	2015.09.01 – 2017.03.31
Name	ODESSA (Online Diarization Enhanced by recent Speaker identification and Sequential learning Approaches)
Funding	SNF – ANR
Coordinator	Centre National de La Recherche Scientifique
Duration	2016.03.01 – 2019.02.28
Partner(s)	Idiap Research Institute, Eurecom
Name	LIVEHEART (The Cellular Basis of Cardiac Development Revealed by Live Imaging)
Funding	SNF – ANR
Coordinator	Institut de Génétique et de Biologie Moléculaire et Cellulaire
Duration	2016.06.01 – 2019.05.31
Partner(s)	Idiap Research Institute, University of Bern, École Polytechnique Paris
Name	TACT-HAND (Improving control of prosthetic hands using tactile sensors and realistic machine learning)
Funding	SNF – DACH
Coordinator	deutsches Zentrum fuer Luft – Und Raumfahrt Ev
Duration	2016.04.01 – 2019.03.31
Partner(s)	Idiap Research Institute, Universitaet Bielefeld
Name	MAAYA (Multimedia Analysis and Access for Documentation and Decipherment of Maya Epigraphy)
Funding	SNF – DACH
Coordinator	Idiap Research Institute
Duration	2013.09.01 – 2017.04.30
Partner(s)	University of Geneva, University of Bonn

Name	COMETS-M (Computational Methods for Temporal Super-resolution Microscopy)
Funding	SNF – Division II
Coordinator	Idiap Research Institute
Duration	2016.04.01 – 2019.03.30
Name	HFACE (HeterogeneousFace Recognition)
Funding	SNF – Division II
Coordinator	Idiap Research Institute
Duration	2014.07.01 – 2018.06.30
Name	ISUL (Importance sampling for large-scale unsupervised learning)
Funding	SNF – Division II
Coordinator	Idiap Research Institute
Duration	2017.03.01 – 2020.02.28
Name	MASS (Multilingual Affective Speech Synthesis)
Funding	SNF – Division II
Coordinator	Idiap Research Institute
Duration	2017.05.01 – 2020.04.30
Name	PHASER-QUAD (Parsimonious Hierarchical Automatic Speech Recognition and Query Detection)
Funding	SNF – Division II
Coordinator	Idiap Research Institute
Duration	2016.10.01 – 2018.09.30
Name	UNITS (Unified Speech Processing Framework for Trustworthy Speaker Recognition)
Funding	SNF – Division II
Coordinator	Idiap Research Institute
Duration	2015.07.01 – 2018.06.30
Name	I-DRESS (Assistive Interactive robotic system for support in DRESSing)
Funding	SNF – ERA NET
Coordinator	Idiap Research Institute
Duration	2015.12.01 – 2018.11.30
Name	PLATFORM-MMD (Platform for Reproducible Acquisition, Processing, and Sharing of Dynamic, Multi-Modal Data)
Funding	SNF – R'equip
Coordinator	Idiap Research Institute
Duration	2016.07.01 – 2017.12.31
Name	WILDTRACK (Tracking in the Wild)
Funding	SNF – Sinergia
Coordinator	Idiap Research Institute
Duration	2014.01.01 – 2017.12.31
Partner(s)	École Polytechnique Fédérale de Lausanne, Eidgenoessische Technische Hochschule Zuerich
Name	MOSPEEDI (Motor Speech Disorders)
Funding	SNF – Sinergia
Coordinator	University of Geneva
Duration	2017.10.01 – 2020.09.30
Partner(s)	Idiap Research Institute, University Hospitals of Geneva, Université Paris 3

ACTIVE AND GRANTED PROJECTS IN 2017

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

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

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

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

xxx

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

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

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

Projects Funded by Industrial Partners

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

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

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

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

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

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

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

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

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

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

ACTIVE AND GRANTED PROJECTS IN 2017

Name	REGENN (Robust Eye-Gaze Estimation Deep Neural Network)
Funding	CTI
Coordinator	Idiap Research Institute
Duration	2017.09.01 – 2018.12.31
Partner(s)	Eyeware
Name	UNICITY (3D scene understanding through machine learning to secure entrance zones)
Funding	CTI
Coordinator	Idiap Research Institute
Duration	2017.03.01 – 2018.11.30
Partner(s)	Hes-So Fribourg, Fastcom Technology SA
Name	IMIM (Intelligent Monitoring for In-line Manufacturing)
Funding	CTI
Coordinator	AISA Automation Industrielle SA
Duration	2016.07.01 – 2018.05.31
Partner(s)	Idiap Research Institute
Name	BIOWAVE (A BIometric Watch Activated by VEins)
Funding	CTI
Coordinator	Centre Suisse D'Electronique et de Microtechnique
Duration	2016.02.01 – 2017.12.31
Partner(s)	Idiap Research Institute, Biowatch SA
Name	3DFINGERVEIN (3D FingerVein Biometrics)
Funding	CTI
Coordinator	Hes-So Valais
Duration	2016.06.01 – 2018.03.31
Partner(s)	Idiap Research Institute, Itservices Sonna Sarl
Name	VIEW-2 (Visibility Improvement for Events Webcasting)
Funding	CTI
Coordinator	Hes-So Valais
Duration	2016.08.01 – 2018.03.31
Partner(s)	Idiap Research Institute, Klewel
Name	FARGO (Convenient and Secure 3D Face Recognition based on RGB-D Cameras)
Funding	CTI
Coordinator	Idiap Research Institute
Duration	2016.05.01 – 2018.02.28
Partner(s)	Keylemon SA
Name	ESGEM (Enhanced Swiss German mEdia Monitoring)
Funding	CTI
Coordinator	Idiap Research Institute
Duration	2016.06.01 – 2018.01.31
Partner(s)	Recapp, Argus der Presse AG
Name	MULTIVEO (High Accuracy Speaker-Independent Multilingual Automatic Speech Recognition System)
Funding	CTI
Coordinator	Idiap Research Institute
Duration	2013.11.01 – 2018.03.31
Partner(s)	Veovox

Projects Awarded in 2017 and Starting in 2018

Name	SHAPED (Speech Hybrid Analytics Platform for consumer and Enterprise Devices)
Funding	CTI
Coordinator	Idiap Research Institute
Duration	2018.03.01 – 2020.02.29
Partner(s)	Logitech S.A.
Name	RISE (Rich Interpersonal Skill analytics for rEcruitment)
Funding	CTI
Coordinator	Idiap Research Institute
Duration	2018.05.01 – 2020.04.30
Partner(s)	University of Lausanne
Name	MEMMO (Memory of Motion)
Funding	H2020-RIA-ICT
Coordinator	Centre national de la recherche scientifique
Duration	2018.01.01 – 2021.12.31
Partner(s)	Idiap Research Institute, University of Edinburgh, Max Planck Society for the Advancement of Sciences, University of Oxford, PAL ROBOTICS SL, AIRBUS SAS, Wandercraft, Centre de médecine physique et de réadaptation, Costain Group PLC
Name	MPM (Multimodal People Monitoring)
Funding	Idiap-CSEM Program
Coordinator	Idiap Research Institute
Duration	2018.02.01 – 2019.01.31
Partner(s)	Centre Suisse d'Electronique et de Microtechnique
Name	RAPPS (Reinforced audio processing via physiological signals)
Funding	Idiap-CSEM Program
Coordinator	Idiap Research Institute
Duration	2018.03.01 – 2019.02.28
Partner(s)	Centre Suisse d'Electronique et de Microtechnique
Name	DOMAT (On-demand Knowledge for Document-level Machine Translation) Funding SNF - Division II
Coordinator	Idiap Research Institute
Duration	2018.01.01 – 2018.12.31
Partner(s)	HES-SO Vaud
Name	ROSALIS (Robot skills acquisition through active learning and social interaction strategies) Funding SNF - Division II
Coordinator	Idiap Research Institute
Duration	2018.04.01 – 2022.03.31
Name	SHISSM (Sparse and hierarchical Structures for Speech Modeling)
Funding	SNF - Division II
Coordinator	Idiap Research Institute
Duration	2018.01.01 – 2021.12.31

Articles in scientific journal 2017

- [1] Erica Madonna, David Ginsbourger and Olivia Martius, A Poisson regression approach to model monthly hail occurrence in Northern Switzerland using large-scale environmental variables, in: Atmospheric Research, 2017
- [2] Marzieh Razavi and Mathew Magimai.-Doss, A Posterior-Based Multi-Stream Formulation for G2P Conversion, in: IEEE Signal Processing Letters, 2017
- [3] Guillaume Heusch, André Anjos and Sébastien Marcel, A reproducible study on remote heart rate measurement, in: arXiv, 2017
- [4] Zeestraten, I. Havoutis, J. Silverio, S. Calinon and D. G. Caldwell, An Approach for Imitation Learning on Riemannian Manifolds, in: IEEE Robotics and Automation Letters (RA-L), 2(3):1240–1247, 2017
- [5] Rui Hu, Carlos Pallan Gayol, Jean-Marc Odobez and Daniel Gatica-Perez, Analyzing and Visualizing Ancient Maya Hieroglyphics Using Shape: from Computer Vision to Digital Humanities, in: Digital Scholarship in the Humanities, 32:179–194, 2017
- [6] Zahid Akhtar, Abdenour Hadid, Mark Nixon, Massimo Tistarelli, Jean-Luc Dugelay and Sébastien Marcel, Biometrics: In Search of Identity and Security (Q & A), in: IEEE MultiMedia, PP, 2017
- [7] Joan-Isaac Biel, Nathalie Martin, David Labbe and Daniel Gatica-Perez, Bites'n'Bits: Inferring Eating Behavior from Contextual Mobile Data, in: Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (PACM IMWUT), 1(4):125–157, 2017
- [8] Milos Cernak, Juan Rafael Orozco-Arroyave, Frank Rudzicz, Heidi Christensen, Juan Camilo Vasquez-Correa and Elmar Nöth, Characterisation of voice quality of Parkinson's disease using differential phonological posterior features, in: Computer Speech and Language, 2017
- [9] Laurent Son Nguyen, Salvador Ruiz-Correa, Marianne Schmid Mast and Daniel Gatica-Perez, Check Out This Place: Inferring Ambiance from Airbnb Photos, in: IEEE transactions on Multimedia, PP(99), 2017
- [10] Milos Cernak, Afsaneh Asaei and Alexandre Hyafil, Cognitive Speech Coding, in: IEEE Signal Processing Magazine, 2017
- [11] N. Jaquier, M. Connan, C. Castellini and S. Calinon, Combining Electromyography and Tactile Myography to Improve Hand and Wrist Activity Detection in Prostheses, in: Technologies, 5(4), 2017
- [12] Amir Mohammadi, Sushil Bhattacharjee and Sébastien Marcel, Deeply Vulnerable -- a study of the robustness of face recognition to presentation attacks, in: IET (The Institution of Engineering and Technology) -- Biometrics:1–13, 2017
- [13] Florian Labhart, Flavio Tarsetti, Olivier Bornet, Darshan Santani, Jasmine Truong, Sara Landolt, Daniel Gatica-Perez and Emmanuel Kuntsche, Development of the Geographical Proportional-to-size Street-Intercept Sampling (GPSIS) method for recruiting urban night-life-goers in an entire city, in: International Journal of Social Research Methodology, 20(6):721–736, 2017
- [14] Kevin G. Chan and Michael Liebling, Direct inversion algorithm for focal plane scanning optical projection tomography, in: Biomedical Optics Express, 2017
- [15] Dario Azzimonti and David Ginsbourger, Estimating orthant probabilities of high dimensional Gaussian vectors with an application to set estimation, in: Journal of Computational and Graphical Statistics, 2017
- [16] Nikolaos Pappas and Andrei Popescu-Belis, Explicit Document Modeling through Weighted Multiple-Instance Learning, in: Journal of Artificial Intelligence Research, 58:591–626, 2017
- [17] Rui Hu, Jean-Marc Odobez and Daniel Gatica-Perez, Extracting Maya Glyphs from Degraded Ancient Documents via Image Segmentation, in: Journal on Computing and Cultural Heritage, 10, 2017
- [18] Pavel Korshunov and Sébastien Marcel, Impact of score fusion on voice biometrics and presentation attack detection in cross-database evaluations, in: IEEE Journal of Selected Topics in Signal Processing, 11(4):695 – 705, 2017
- [19] E. Pignat and S. Calinon, Learning adaptive dressing assistance from human demonstration, in: Robotics and Autonomous Systems, 93:61–75, 2017
- [20] D. Bruno, S. Calinon and D. G. Caldwell, Learning Autonomous Behaviours for the Body of a Flexible Surgical Robot, in: Autonomous Robots, 41(2):333–347, 2017
- [21] Hannah Muckenhirn, Pavel Korshunov, Mathew Magimai.-Doss and Sébastien Marcel, Long-Term Spectral Statistics for Voice Presentation Attack Detection, in: IEEE/ACM Transactions on Audio, Speech and Language Processing, 25(11):2098–2111, 2017
- [22] Riwal Lefort, L. Fusco, O. Pertz and Francois Fleuret, Machine learning-based tools to model and to remove the off-target effect, in: Pattern Analysis and Applications, 20(1):87–100, 2017

Book and book chapters 2017

- [23] Nikolaos Pappas, Miriam Redi, Mercan Topkara, Hongyi Liu, Brendan Jou, Tao Chen and Shih-Fu Chang, Multilingual Visual Sentiment Concept Clustering and Analysis, in: *International Journal of Multimedia Information Retrieval*, 2017
- [24] Juan Rafael Orozco-Arroyave, Juan Camilo Vasquez-Correa, Jesús Francisco Vargas-Bonilla, Raman Arora, Najim Dehak, Phani Sankar Nidadavolu, Heidi Christensen, Frank Rudzicz, Maria Yancheva, Alyssa Vann, Nikolai Vogler, Tobias Bocklet, Milos Cernak, Julius Hannink and Elmar Nöth, NeuroSpeech: An open-source software for Parkinson's speech analysis, in: *Digital Signal Processing*, 2017
- [25] Niklas Linde, David Ginsbourger, James Irving, Fabio Nobile and Arnaud Doucet, On uncertainty quantification in hydrogeology and hydrogeophysics, in: *Advances in Water Resources*, 110:166–181, 2017
- [26] Afsaneh Aseai, Milos Cernak and Hervé Bourlard, Perceptual Information Loss due to Impaired Speech Production, in: *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2017
- xxxiv
- [27] Aleksandra Cerekovic, Oya Aran and Daniel Gatica-Perez, Rapport with Virtual Agents: What do Human Social Cues and Personality Explain?, in: *IEEE Transactions on Affective Computing*, 8(3):382–395, 2017
- [28] Salvador Ruiz-Correa, Darshan Santani, Beatriz Ramirez Salazar, Itzia Ruiz-Correa, Fatima Rendon-Huerta, Carlo Olmos Carrillo, Brisa Sandoval Mexicano, Angel Arcos Garcia, Rogelio Hasimoto-Beltran and Daniel Gatica-Perez, SenseCityVity: Mobile Crowdsourcing, Urban Awareness, and Collective Action in Mexico, in: *IEEE Pervasive Computing*, Special Issue on Smart Cities, 16(2):44–53, 2017
- [29] Subhadeep Dey, Petr Motlicek, Srikanth Madikeri and Marc Ferras, Template-matching for Text-dependent Speaker Verification, in: *Speech Communication*, 2017
- [30] R. Reiter-Palmon, T. Sinha, J. Gevers, Jean-Marc Odobez and G. Volpe, Theories and Models of Teams and Group, in: *Small Group Research*, 48(5):544–567, 2017
- [1] Daniel Gatica-Perez, Oya Aran and Dinesh Babu Jayagopi, Analysis of Small Groups, in: *Social Signal Processing*, pages 349–367, Cambridge University Press. Editors J. Burgoon, N. Magnenat-Thalmann, M. Pantic, and A. Vinciarelli, 2017
- [2] Daniel Gatica-Perez, Gulcan Can, Rui Hu, Stephane Marchand-Maillet, Jean-Marc Odobez, Carlos Pallan Gayol and Edgar Roman-Rangel, MAAAYA: Multimedia Methods to Support Maya Epigraphic Analysis, in: *Arqueologia computacional: Nuevos enfoques para el analisis y la difusion del patrimonio cultural*, INAH-RedTDPC, 2017
- [3] Pavel Korshunov and Sébastien Marcel, Presentation attack detection in voice biometrics, in: *User-Centric Privacy and Security in Biometrics*, The Institution of Engineering and Technology, 2017

Articles in conferences proceedings 2017

- [1] Trung Phan Thanh and Daniel Gatica-Perez, #Healthy #Fondue #Dinner: Analysis and Inference of Food and Drink Consumption Patterns on Instagram, in: Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia, 2017
- [2] Z. Boulkenafet, J. Komulainen, Z. Akhtar, A. Benlamoudi, SE. Bekhouche, F. Dornaika, A. Ouafi, Amir Mohammadi, Sushil Bhattacharjee and Sébastien Marcel, A Competition on Generalized Software-based Face Presentation Attack Detection in Mobile Scenarios, in: Proceedings of the International Joint Conference on Biometrics, 2017, 2017
- [3] Youssef Oualil, Dietrich Klakow, Gyorgy Szaszak, Ajay Srinivasamurthy, Hartmut Helmke and Petr Motlicek, A Context-Aware Speech recognition and Understanding System for Air Traffic Control Domain, in: Proceedings of the IEEE Automatic Speech Recognition and Understanding Workshop, Okinawa, Japan, 2017
- [4] Nam Le and Jean-Marc Odobez, A Domain Adaptation Approach to Improve Speaker Turn Embedding Using Face Representation, in: ACM International Conference on Multimodal Interaction, Glasgow, Scotland, ACM, 2017
- [5] Ajay Kumar Tanwani and S. Calinon, A Generative Model for Intention Recognition and Manipulation Assistance in Teleoperation, in: IEEE/RSJ International Conference on Intelligent Robots and Systems, 2017
- [6] James Newling and Francois Fleuret, A Sub-Quadratic Exact Medoid Algorithm, in: Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, 2017
- [7] Jagannadan Varadarajan, R Subramanian, Narendra Ahuja, Pierre Moulin and Jean-Marc Odobez, Active Online Anomaly Detection using Dirichlet Process Mixture Model and Gaussian Process Classification, in: IEEE Winter Conference on Applications of Computer Vision (WACV), Washington, 2017
- [8] Sibong Tong, Philip N. Garner and Hervé Boudlard, An Investigation of Deep Neural Networks for Multilingual Speech Recognition Training and Adaptation, in: Proc. of Interspeech, 2017
- [9] André Anjos, Laurent El Shafey and Sébastien Marcel, BEAT: An Open-Science Web Platform, in: Thirty-fourth International Conference on Machine Learning, Sydney, Australia, 2017
- [10] Milos Cernak, Alain Komaty, Amir Mohammadi, André Anjos and Sébastien Marcel, Bob Speaks Kaldi, in: Proc. of Interspeech, 2017
- [11] Samaneh Abbasi-Sureshjani, Behdad Dasht Bozorg, Bart ter Haar Romeny and Francois Fleuret, Boosted Exudate Segmentation in Retinal Images using Residual Nets, in: Proceedings of the MICCAI Workshop on Ophthalmic Medical Image Analysis, 2017
- [12] Xiao Pu, Laura Mascarell and Andrei Popescu-Belis, Consistent Translation of Repeated Nouns using Syntactic and Semantic Cues, in: Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL), Valencia, Spain, Association for Computational Linguistics, 2017
- [13] Subhadeep Dey, Srikanth Madikeri, Petr Motlicek and Marc Ferras, Content Normalization for Text-dependent Speaker Verification, in: Proc. of Interspeech, 2017
- [14] André Anjos, Manuel Günther, Tiago de Freitas Pereira, Pavel Korshunov, Amir Mohammadi and Sébastien Marcel, Continuously Reproducing Toolchains in Pattern Recognition and Machine Learning Experiments, in: Thirty-fourth International Conference on Machine Learning, Sidney, Australia, 2017
- [15] Ana Sequeira, Lulu Chen, James Ferryman, Peter Wild, Fernando Alonso-Fernandez, Josef Bigün, Kiran B. Raja, R. Raghavendra, Christoph Busch, Tiago de Freitas Pereira, Sébastien Marcel, Sushree Sangeeta Behera, Mahesh Gour and Vivek Kanhangad, Cross-Eyed 2017: Cross-Spectral Iris/Periocular Recognition Competition., in: IEEE/IAPR International Joint Conference on Biometrics, Denver, Colorado, USA, IEEE, 2017
- [16] Tatjana Chavdarova and Francois Fleuret, Deep Multi-Camera People Detection, in: Proceedings of the IEEE International Conference on Machine Learning and Applications, 2017
- [17] Pierre Baqué, Francois Fleuret and Pascal Fua, Deep Occlusion Reasoning for Multi-Camera Multi-Target Detection, in: Proceedings of the IEEE International Conference on Computer Vision, 2017
- [18] D. Berio, S. Calinon and F. F. Leymarie, Dynamic Graffiti Stylisation with Stochastic Optimal Control, in: Intl Workshop on movement and computing (MOCO), London, UK, pages 1–8, ACM, 2017
- [19] Rui Hu, Hieu Pham, Philipp Bulushek and Daniel Gatica-Perez, Elderly People Living Alone: Detecting Home Visits with Ambient and Wearable Sensing, in: Proceedings of MMHealth, 2017
- [20] Hannah Muckenhirn, Mathew Magimai.-Doss and Sébastien Marcel, End-to-End Convolutional Neural Network-based Voice Presentation Attack Detection, in: International Joint Conference on Biometrics, Denver, Colorado, USA, 2017

LIST OF PUBLICATIONS IN 2017

- [21] Skanda Muralidhar and Daniel Gatica-Perez, Examining Linguistic Content and Skill Impression Structure for Job Interview Analytics in Hospitality, in: Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia, 2017
- [22] Pranay Dighe, Afsaneh Asaei and Hervé Bourlard, Exploiting Eigenposteriors for Semi-supervised Training of DNN Acoustic Models with Sequence Discrimination, in: Proceedings of Interspeech, 2017
- [23] Subhadeep Dey, Petr Motlicek, Srikanth Madikeri and Marc Ferras, EXPLOITING SEQUENCE INFORMATION FOR TEXT-DEPENDENT SPEAKER VERIFICATION, in: Proceedings of 2017 IEEE International Conference on Acoustics, Speech, and Signal Processing, New Orleans, pages 5370–5374, 2017
- [24] Samaneh Abbasi-Sureshjani, Behdad Dasht Bozorg, Bart ter Haar Romeny and Francois Fleuret, Exploratory Study on Direct Prediction of Diabetes using Deep Residual Networks, in: Proceedings of the thematic conference on computational vision and medical image processing, 2017
- [25] N. Jaquier and S. Calinon, Gaussian Mixture Regression on Symmetric Positive Definite Matrices Manifolds: Application to Wrist Motion Estimation with sEMG, in: Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems, 2017
- [26] D. Berio, S. Calinon and F. F. Leymarie, Generating Calligraphic Trajectories with Model Predictive Control, in: Proc. 43rd Conf. on Graphics Interface, Edmonton, AL, Canada, pages 132–139, 2017
- [27] Skanda Muralidhar, Marianne Schmid Mast and Daniel Gatica-Perez, How May I Help You? Behavior and Impressions in Hospitality Service Encounters, in: Proceedings of 19th ACM International Conference on Multimodal Interaction, 2017
- [28] N. Jaquier, C. Castellini and S. Calinon, Improving hand and wrist activity detection using tactile sensors and tensor regression methods on Riemannian manifolds, in: Proc. of the Myoelectric Control Symposium, 2017
- [29] Nam Le and Jean-Marc Odobez, Improving speaker turn embedding by crossmodal transfer learning from face embedding, in: ICCV Workshop on Computer Vision for Audio-Visual Media, 2017
- [30] Darshan Santani, Salvador Ruiz-Correa and Daniel Gatica-Perez, Insiders and Outsiders: Comparing Urban Impressions between Population Groups, in: International Conference on Multimedia Retrieval, ACM, 2017
- [31] Srikanth Madikeri, Marc Ferras, Petr Motlicek and Subhadeep Dey, INTRA-CLASS COVARIANCE ADAPTATION IN PLDA BACK-ENDS FOR SPEAKER VERIFICATION, in: Proceedings of International Conference on Acoustics, Speech and Signal Processing, pages 5365–5369, 2017
- [32] James Newling and Francois Fleuret, K-Medoids For K-Means Seeding, in: Proceedings of the international conference on Neural Information Processing Systems, 2017
- [33] L. Rozo, N. Jaquier, S. Calinon and D. G. Caldwell, Learning Manipulability Ellipsoids for Task Compatibility in Robot Manipulation, in: Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems, 2017
- [34] M. Zeestraten, I. Havoutis, S. Calinon and D. G. Caldwell, Learning Task-Space Synergies using Riemannian Geometry, in: Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS), Vancouver, Canada, pages 73–78, IEEE, 2017
- [35] Pranay Dighe, Afsaneh Asaei and Hervé Bourlard, Low-rank and Sparse Soft Targets to Learn Better DNN Acoustic Models, in: Proceedings of 2017 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2017), 2017
- [36] Matthias Kleinert, Hartmut Helmke, Gerald Siol, heiko Ehr, Michael Finke, Youssef Oualil and Ajay Srinivasamurthy, Machine Learning of Controller Command Prediction Models from Recorded Radar Data and Controller Speech Utterances, in: Proceedings of the 7th SESAR Innovation Days (SID), University of Belgrade, Belgrade, Serbia, 2017
- [37] Ngoc-Quang Luong and Andrei Popescu-Belis, Machine translation of Spanish personal and possessive pronouns using anaphora probabilities, in: Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL), Valencia, Spain, Association for Computational Linguistics, 2017
- [38] Pierre Baqué, Francois Fleuret and Pascal Fua, Multi-Modal Mean-Fields via Cardinality-Based Clamping, in: Proceedings of the IEEE international conference on Computer Vision and Pattern Recognition, 2017
- [39] Juan Camilo Vasquez-Correa, Juan Rafael Orozco-Arroyave, Raman Arora, Elmar Nöth, Najim Dehak, Heidi Christensen, Frank Rudzicz, Tobias Bocklet, Milos Cernak, Hamidreza Chinaei, Julius Hannink, Phani Sankar Nidadavolu, Maria Yancheva, Alyssa Vann and Nikolai Vogler, Multi-view Representation Learning Via GCCA for Multimodal Analysis of Parkinson's Disease, in: Proceedings of 2017 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2017), 2017
- [40] Nikolaos Pappas and Andrei Popescu-Belis, Multilingual Hierarchical Attention Networks for Document Classification, in: 8th International Joint Conference on Natural Language Processing (IJCNLP), pages 1015–1025, 2017
- [41] Andrii Maksai, Xinchao Wang, Francois Fleuret and Pascal Fua, Non-Markovian Globally Consistent Multi-Object Tracking, in: Proceedings of the IEEE International Conference on Computer Vision, 2017

LIST OF PUBLICATIONS IN 2017

- [42] Sébastien Marmin, Jean Baccou, Jacques Liandrat and David Ginsbourger, Non-parametric warping via local scale estimation for non-stationary Gaussian process modelling, in: *Wavelets and Sparsity XVII*, pages 1039-1042, International Society for Optics and Photonics, 2017
- [43] Skanda Muralidhar, On Job Training: Automated Interpersonal Behavior Assessment & Real-Time Feedback, in: *Proceedings of the 25th ACM International Conference on Multimedia*, 2017, 2017
- [44] Andréé R. Goncalves, Pavel Korshunov, Ricardo P. V. Violato, Flávio O. Simões and Sébastien Marcel, On the Generalization of Fused Systems in Voice Presentation Attack Detection, in: *16th International Conference of the Biometrics Special Interest Group*, 2017
- [45] Milos Cernak, Elmar Nöth, Frank Rudzicz, Heidi Christensen, Juan Rafael Orozco-Arroyave, Raman Arora, Tobias Bocklet, Hamidreza Chinaei, Julius Hannink, Phani Sankar Nidadavolu, Juan Camilo Vasquez, Maria Yancheva, Alyssa Vann and Nikolai Vogler, On the Impact of Non-modal Phonation On Phonological Features, in: *Proceedings of 2017 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2017)*, 2017
- [46] Sébastien Marmin, Jean Baccou, Frédéric Perales, David Ginsbourger and Jacques Liandrat, Planification adaptative d'expériences numériques par paquets en contexte non stationnaire pour une étude de fissuration mécanique, in: *23eme Congres Francais de Mecanique*, 28 aout – 1er septembre 2017, Lille, France (FR), AFM, 2017
- [47] Yu Yu, Kenneth Alberto Funes Mora and Jean-Marc Odobez, Robust and Accurate 3D Head Pose Estimation through 3DMM and Online Head Model Reconstruction, in: *Proceedings of the 12th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2017)*, 2017
- [48] Ajay Srinivasamurthy, Petr Motlicek, Ivan Himawan, Gyorgy Szaszak, Youssef Oualil and Hartmut Helmke, Semi-supervised Learning with Semantic Knowledge Extraction for Improved Speech Recognition in Air Traffic Control, in: *Proceedings of Interspeech 2017*, Stockholm, Sweden, pages 2406–2410, 2017
- [49] Xiao Pu, Nikolaos Pappas and Andrei Popescu-Belis, Sense-Aware Statistical Machine Translation using Adaptive Context-Dependent Clustering, in: *Proceedings of Second Conference on Machine Translation (WMT17)*, 2017
- [50] Gulcan Can, Jean-Marc Odobez and Daniel Gatica-Perez, Shape Representations for Maya Codical Glyphs: Knowledge-driven or Deep?, in: *15th International Workshop on Content-Based Multimedia Indexing*, 2017
- [51] Timur Bagautdinov, Alexandre Alahi, Francois Fleuret, Pascal Fua and Sylvio Savarese, Social Scene Understanding: End-to-End Multi-Person Action Localization and Collective Activity Recognition, in: *Proceedings of the IEEE international conference on Computer Vision and Pattern Recognition*, 2017
- [52] Afsaneh Asaei, Milos Cernak, Hervé Bouldard and Dhananjay Ram, Sparse Pronunciation Codes for Perceptual Phonetic Information Assessment, in: *Workshop on Signal Processing with Adaptive Sparse Structured Representations (SPARS)*, 2017
- [53] Dhananjay Ram, Afsaneh Asaei and Hervé Bouldard, Subspace Regularized Dynamic Time Warping for Spoken Query Detection, in: *Workshop on Signal Processing with Adaptive Sparse Structured Representations (SPARS)*, 2017
- [54] I. Havoutis and S. Calinon, Supervisory teleoperation with online learning and optimal control, in: *Proc. IEEE Intl Conf. on Robotics and Automation (ICRA)*, Singapore, pages 1534–1540, IEEE, 2017
- [55] Renars Liepins and et al., The SUMMA Platform Prototype, in: *Proceedings of the EACL 2017 Software Demonstrations*, Valencia, Spain, pages 116–119, 2017
- [56] Khaled Khelif, Yann Mombrun, Gerhard Backfried, Farhan Sahito, Luca Scarpato, Petr Motlicek, Damien Kelly, Gideon Hazzani, Emmanouil Chatzigavriil and Srikanth Madikeri, Towards a breakthrough Speaker Identification approach for Law Enforcement Agencies: SIIP, in: *European Intelligence and Security Informatics Conference (EISIC) 2017*, Athenes, Greece, pages 32–39, IEEE Computer Society, 2017
- [57] Nam Le, Jean-Marc Odobez and et al., Towards large scale multimedia indexing: A case study on person discovery in broadcast news, in: *15th International Workshop on Content-Based Multimedia Indexing*, 2017
- [58] Remy Siegfried, Yu Yu and Jean-Marc Odobez, Towards the Use of Social Interaction Conventions as Prior for Gaze Model Adaptation, in: *Proceedings of 19th ACM International Conference on Multimodal Interaction*, pages 9, ACM, 2017
- [59] C. Mastalli, M. Focchi, I. Havoutis, A. Radulescu, S. Calinon, J. Buchli, D. G. Caldwell and C. Semini, Trajectory and Foothold Optimization using Low-Dimensional Models for Rough Terrain Locomotion, in: *Proc. IEEE Intl Conf. on Robotics and Automation (ICRA)*, Singapore, pages 1096–1103, IEEE, 2017
- [60] Lesly Miculicich Werlen and Andrei Popescu-Belis, Using Coreference Links to Improve Spanish-to-English Machine Translation, in: *Proceedings of the 2nd Workshop on Coreference Resolution Beyond OntoNotes (CORBON 2017)*, Valencia, Spain, pages 30–40, Association for Computational Linguistics, 2017

PHD THESES 2017

- [61] Lesly Miculicich Werlen and Andrei Popescu-Belis, Validation of an Automatic Metric for the Accuracy of Pronoun Translation (APT), in: Proceedings of the Third Workshop on Discourse in Machine Translation (DiscoMT), Denmark, Copenhagen, Association for Computational Linguistics (ACL), 2017
- [62] Yassir Benkhedda, Darshan Santani and Daniel Gatica-Perez, Venues in Social Media: Examining Ambiance Perception Through Scene Semantics, in: Proceedings of the 25th ACM International Conference on Multimedia, ACM, 2017, 2017
- [63] Stepan Tulyakov, Anton Ivanov and Francois Fleuret, Weakly Supervised Learning of Deep Metrics for Stereo Reconstruction, in: Proceedings of the IEEE International Conference on Computer Vision, 2017
- [64] Sushil Bhattacharjee and Sébastien Marcel, What you can't see can help you -- extended-range imaging for 3D-mask presentation attack detection, in: Proceedings of the 16th International Conference on Biometrics Special Interest Group., Darmstadt (Germany), Gesellschaft fuer Informatik e.V. (GI), 2017
- [1] Pierre-Edouard Honnet, Intonation Modelling for Speech Synthesis and Emphasis Preservation, École Polytechnique Fédérale de Lausanne, 2017
- [2] Pedro H. O. Pinheiro, Large-Scale Image Segmentation with Convolutional Networks, Sciences et Techniques de l'Ingénieur (STI), 2017
- [3] Olivier Canévet, Object Detection with Active Sample Harvesting, École Polytechnique Fédérale de Lausanne, 2017
- [4] Marzieh Razavi, On Modeling the Synergy Between Acoustic and Lexical Information for Pronunciation Lexicon Development, École polytechnique fédérale de Lausanne (EPFL), 2017
- [5] Gulcan Can, Visual Analysis of Maya Glyphs via Crowdsourcing and Deep Learning, École Polytechnique Fédérale de Lausanne, 2017