idiap
INSTITUT DE RECHERCHE

RAPPORT ANNUEL 2016

réalisation Martina Fellay, Christine Marcel et Léonore Miauton, équipe gestion de projets
rédaction Philippe Morel **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.loterie.ch

www.swisscom.ch

www.groupemutuel.ch

www.epfl.ch

www.theark.ch

www.ideark.ch

www.snf.ch

www.kti.admin.ch

www.haslerstiftung.ch

SOMMAIRE

Messages

04 Un anniversaire entre succès et défis
05 Retour aux origines 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
16 Collaborateurs

Finances

19 Compte d'exploitation
20 Sources de financement
21 Bilan

Recherche

23 L'enseignement à distance 3.0
24 L'Idiap maintient le cap !
25 La biométrie au service de la cybersécurité
26 Baxter, le robot qui vous habille
27 De prestigieuses collaborations avec les USA
28 Conférence internationale en « deep learning »
29 Moi, Pepper, robot humanoïde

Réseau

31 Le retour au Pays
32 L'Idiap fête ses 25 ans !
34 ICC'2016 : de l'idée à la start-up
36 L'Idiap à la rencontre des jeunes

Visages

39 Dribbler avec les chiffres
40 Valais-Wallis Ambition : 3 étudiants du programme
42 Distinctions
43 Thèses achevées

Encarté scientifique

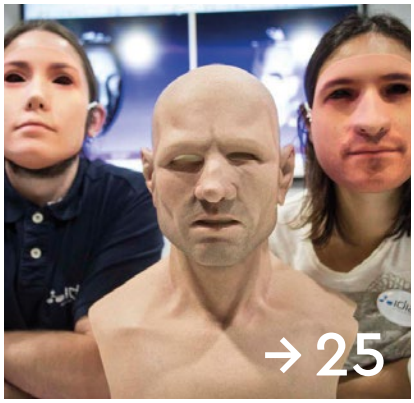
II Scientific Report
XXVII Main projects in progress
XXXV Major publications / Conferences

→ 24

→ 29

→ 31

→ 33


→ 25

→ 28

Un anniversaire entre succès et défis

Sobrement, mais non sans fierté, l'Idiap a fêté ses 25 ans en 2016. De nombreux événements ont célébré, tout au long de l'année, ce quart de siècle (pp. 32-33).

L'Institut jouit désormais d'une indéniable reconnaissance dans le monde de l'intelligence artificielle. Cet enviable statut est en grande partie dû aux compétences scientifiques et qualités d'animateur, de gestionnaire et de visionnaire de son directeur, le Professeur Hervé Bourlard. Motivant ses troupes, créatif et exigeant, il a su en 20 ans de direction mixer les ingrédients qui permettent aujourd'hui à l'Idiap d'initier d'actives collaborations avec les plus grands : Google, Facebook, Samsung, etc. Et ce, bien évidemment, sans oublier les relations académiques nouées aux quatre coins de la planète. Toutefois, cette magnifique « success story » n'est pas un aboutissement mais une plateforme sur laquelle s'appuyer pour poursuivre nos recherches faites d'opportunités et d'innovations répondant aux besoins de la société. Nous sommes entrés dans l'ère que d'aucuns appellent la quatrième révolution industrielle. Elle se caractérise, peu ou prou, par la digitalisation du savoir et sa large diffusion, par la numérisation des services et par la communication entre les objets sans intervention directe de l'humain.

Les effets économiques et philosophiques qu'entraînera ce nouveau monde hyper-technologique restent difficilement imaginables. Toutefois, quelques interrogations sociétales sont d'ores et déjà perceptibles. Ainsi, par exemple, chacun salue le déploiement des formations permettant de disposer de scientifiques et techniciens capables de concevoir et installer les robots qui accomplissent les tâches confiées autrefois à un opérateur humain. Cependant, n'est-il pas légitime d'engager un débat à dessein de garantir un revenu minimal à ceux qui se retrouveraient sans travail ? Ou encore, comment considérer les innovations du secteur de la santé, riches de grandes promesses mais de périls aussi ? S'il est extraordinaire d'imaginer vivre en pleine forme jusqu'à 100 ans et bénéficier d'un véhicule autonome pour se rendre chez son médecin, de profondes questions subsistent sur l'acceptation de cet « homme amélioré » dans « une société digitalisée ». C'est dans ce monde en rapide évolution que doivent se mouvoir les collaborateurs de l'Idiap. Leur dynamisme et leurs larges compétences nous permettront de participer et résister à d'inévitables bouleversements. Institut de recherche et centre de compétences innovant dans les sciences informatiques, en particulier dans les interactions entre l'homme et la machine, l'Idiap doit cependant renforcer ses collaborations et préciser sa position stratégique dans le paysage suisse de la recherche et de l'innovation. Il importe également de développer la contribution au transfert de technologies dans le secteur industriel, en collaboration avec le partenaire privilégié qu'est l'EPFL et avec les universités cantonales et les Hautes écoles spécialisées.

C'est ainsi que l'Idiap filera sereinement vers son 50e anniversaire !

« L'Idiap initie d'actives collaborations avec les plus grands »

Olivier Dumas
Président du Conseil de fondation de l'Idiap

Retour aux origines de l'Idiap

« Intelligence artificielle : retour aux sources de l'Idiap »

Hervé Bourlard
Directeur de l'Idiap, Professeur EPFL

Depuis quelques temps, l'intelligence artificielle fait les titres de la presse quotidienne et généraliste, bien souvent en lien avec la quatrième révolution industrielle. Sous ce terme, usité de manière assez large, se cachent également les vocables de « deep learning » et de « machine learning ». En tant que directeur de l'Idiap, je ne peux que me réjouir que le formidable pari entrepris il y a 25 ans donne aujourd'hui raison aux visionnaires qui ont fondé l'Idiap et aux autorités qui les ont soutenus.

Lorsque je suis arrivé à Martigny, en 1996, mon but était de créer un institut multidisciplinaire de réputation internationale, actif dans les domaines du traitement du signal et de l'apprentissage automatique, comme technologies communes aux domaines d'application du traitement (reconnaissance) de la parole et de la vision par ordinateur. A une époque où la multidisciplinarité n'était pas encore une évidence, voire même une obligation, on me laissait la liberté de viser cet objectif, de tisser des liens uniques entre ces disciplines importantes et d'y intégrer, au fil des ans, d'autres domaines de recherche, comme l'analyse multi-senseurs, le traitement de l'information multimédia, et la biométrie, par exemple.

A partir de financements assez limités, et dans le monde ultra-compétitif de la science, où seule la réputation d'une institution attire les meilleurs chercheurs, développer un nouvel institut de recherche de très haut niveau n'a pas été chose aisée. Les scientifiques de l'Idiap ont dû s'affirmer et se battre pour bâtir notre réputation et faire jouer notre institut dans la cour des grands. L'attribution par la Confédération du Pôle de recherche national IM2 à l'Idiap, en 2001, a fortement contribué à ce succès.

En 1991, l'Idiap s'appelait officiellement Institut Dalle Molle d'Intelligence Artificielle Perceptive (IDIAP) - aujourd'hui, je parlerais plutôt d'intelligence artificielle perceptuelle (ou encore cognitive). A cette époque, la notion d'intelligence artificielle était au creux de la vague, au profit de l'émergence de nouvelles méthodologies d'apprentissage automatique, basées sur des modèles statistiques, dont notamment les réseaux de neurones si populaires aujourd'hui (souvent associés aux noms de « deep learning » ou encore « deep neural networks »). C'est d'ailleurs la raison pour laquelle le nom de l'Institut de recherche Idiap s'est peu à peu substitué à son acronyme d'origine IDIAP. Ironiquement, c'est maintenant que ces cinq lettres seules se sont imposées que l'intelligence artificielle revient sous les feux de la rampe. Je remercie tous les collaborateurs de l'Idiap pour avoir contribué activement à ce come-back et espère que nous pourrons continuer à construire sur nos acquis, ainsi que notre culture multidisciplinaire.

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-cinq ans plus tard, l'institut place toujours les intérêts de la société au cœur de sa démarche.

Déferlement d'outils technologiques

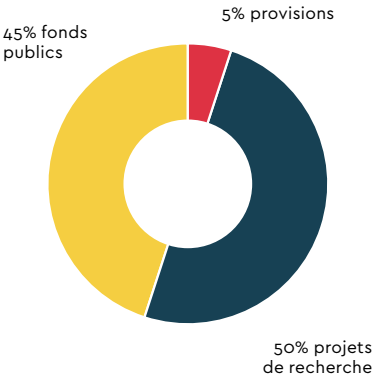
En ce début de XXIe siècle, on assiste à un déferlement permanent de nouveaux outils technologiques. S'ils permettent des gains considérables en termes de productivité et de confort, ils bouleversent aussi les habitudes des gens, laissant une part des utilisateurs démunis et une autre lassée par la modification récurrente des systèmes. Dans ce contexte, l'Idiap travaille essentiellement à l'amélioration des interactions 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.

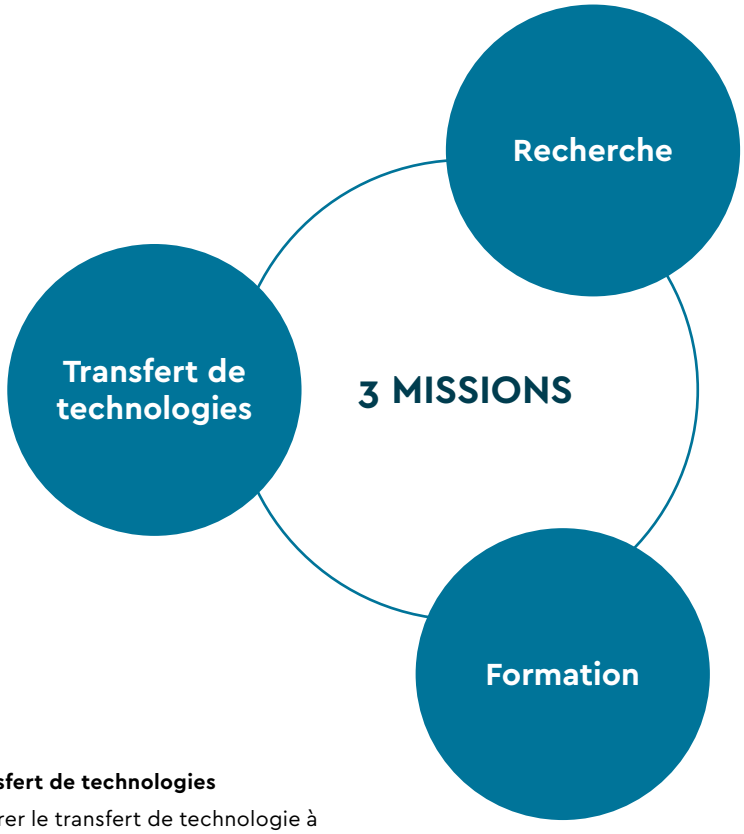
Structure

- Fondation à but non lucratif
- Institution indépendante (mais liée à l'EPFL par un plan de développement commun)
- Spécialisée dans 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



Les trois 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 de l'innovation » (art.15). Sa vision est de promouvoir la qualité de la vie par le progrès scientifique en lien avec le traitement de l'information multimédia et sensorielle. Les activités de l'Idiap s'orientent selon 3 missions complémentaires :



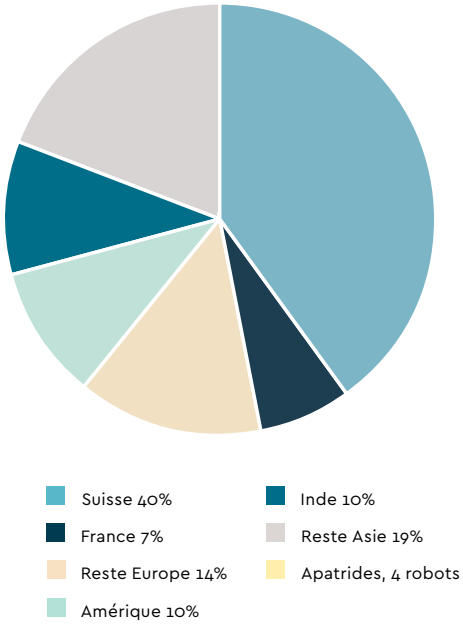
Recherche
Mener des projets de recherche fondamentale au plus haut niveau dans ses domaines de prédilection, s'assurant ainsi une place parmi les meilleurs à l'échelle nationale, européenne et mondiale. Sur la scène internationale, l'Idiap bénéficie d'un large réseau de partenaires et collabore activement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

L'Idiap en chiffres (2016)

AU TOTAL, 25 NATIONALITÉS SONT REPRÉSENTÉES AU SEIN DE L'IDIAP



Ressources humaines

2 professeurs
2 maîtres d'enseignement et de recherche (MER)
8 chercheurs permanents
6 chercheurs associés
18 postdoctorants
31 doctorants
11 ingénieurs de développement
6 ingénieurs système
7 collaborateurs administratifs
11 stagiaires (moyenne/année)

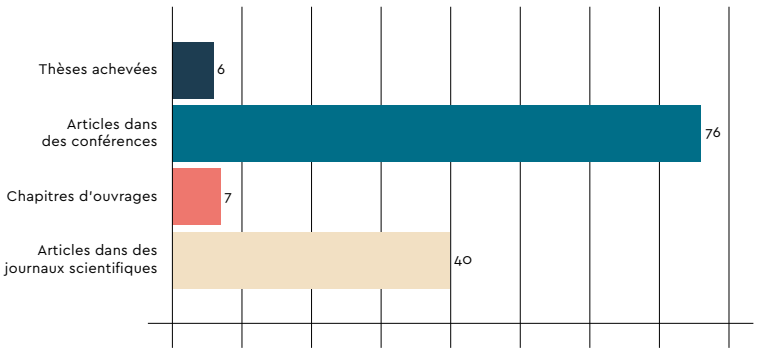
136 personnes au total
94 équivalents plein temps

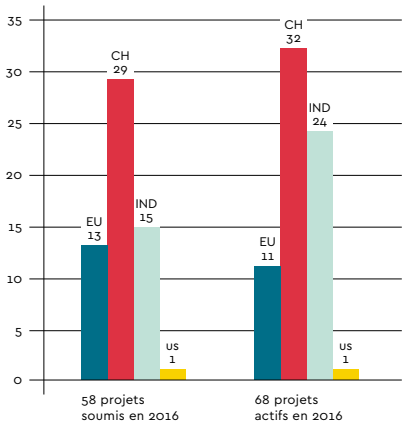
Activités scientifiques

- Participation à 68 programmes de recherche
- Direction de projets dans plus de 48 consortiums (Suisse, Europe, projets industriels)
- Contribution à la stratégie de développement économique du Canton du Valais à travers le programme The Ark et en particulier l'incubateur IdeArk à Martigny
- Plus de 50 postes dans les start-ups IdeArk

Soumission et financement des projets de recherche

En 2016, l'Idiap a soumis 58 projets (29 à une agence de financement suisse, 13 à un fonds européen, 15 projets industriels et 1 projet américain). En parallèle, l'institut était actif dans 68 projets.





Publications

En 2016, les scientifiques de l'Idiap ont participé à 129 publications revues par les pairs.

5 thèmes de recherche
10 domaines d'application

Systèmes perceptifs et cognitifs
Reconnaissance automatique de la parole, vision par ordinateur, reconnaissance de l'écriture, traitement de documents multimédias, robotique, traduction automatique.

Modélisation du comportement humain
Médias de communication web et mobiles, interactions sociales, traitement des signaux sociaux.

Interfaces d'information et de présentation
Systèmes d'information multimédias, contextualisation et personnalisation.

Cyber-sécurité
Authentification de personnes, vérification des caractéristiques biométriques (voix, visage, veines), détection d'activités criminelles.

Apprentissage automatique
Modélisation statistique, réseaux de neurones, modèles mathématiques.

Interaction homme-homme et homme-machine/robot

Commande d'appareils par la voix, systèmes de traduction voix/voix, systèmes de navigation, synthèse de gestes, apprentissages par démonstration.

Exploitation d'archives multimédias riches

Indexation sémantique, reconnaissance d'objets, résumé de contenu audiovisuel, analyse de documents YouTube, analyse de l'héritage culturel.

Systèmes collaboratifs et créatifs

Salle de réunion intelligente, vidéo-conférence, accès et indexation multimédias, collaboration interlangues.

Gestion des systèmes mobiles

Traitement du signal pour plateformes mobiles, recommandation et publicité sur mobile.

Gestion de l'énergie

Réseaux énergétiques hétérogènes, réseaux de senseurs, modélisation de l'activité humaine pour anticiper les besoins.

Santé et bio-engineering

Gestion intelligente des patients, prothèses (auditives, corporelles, etc.), modélisation de systèmes biologiques, interfaces pour personnes en situation de handicap, acquisition et indexation de documents médicaux, bio-imagerie, microscopie.

Sécurité et gestion des risques

Contrôles d'accès (espaces physiques ou virtuels), vérification du locuteur, vérification du visage, vidéosurveillance, modélisation des risques naturels, génie industriel, sécurité des espaces publics et privés.

Villes du futur (« Smart Cities »)

Ecologie, management environnemental, réduction de la pollution, réduction du trafic, meilleur usage du réseau routier, réduction du bruit.

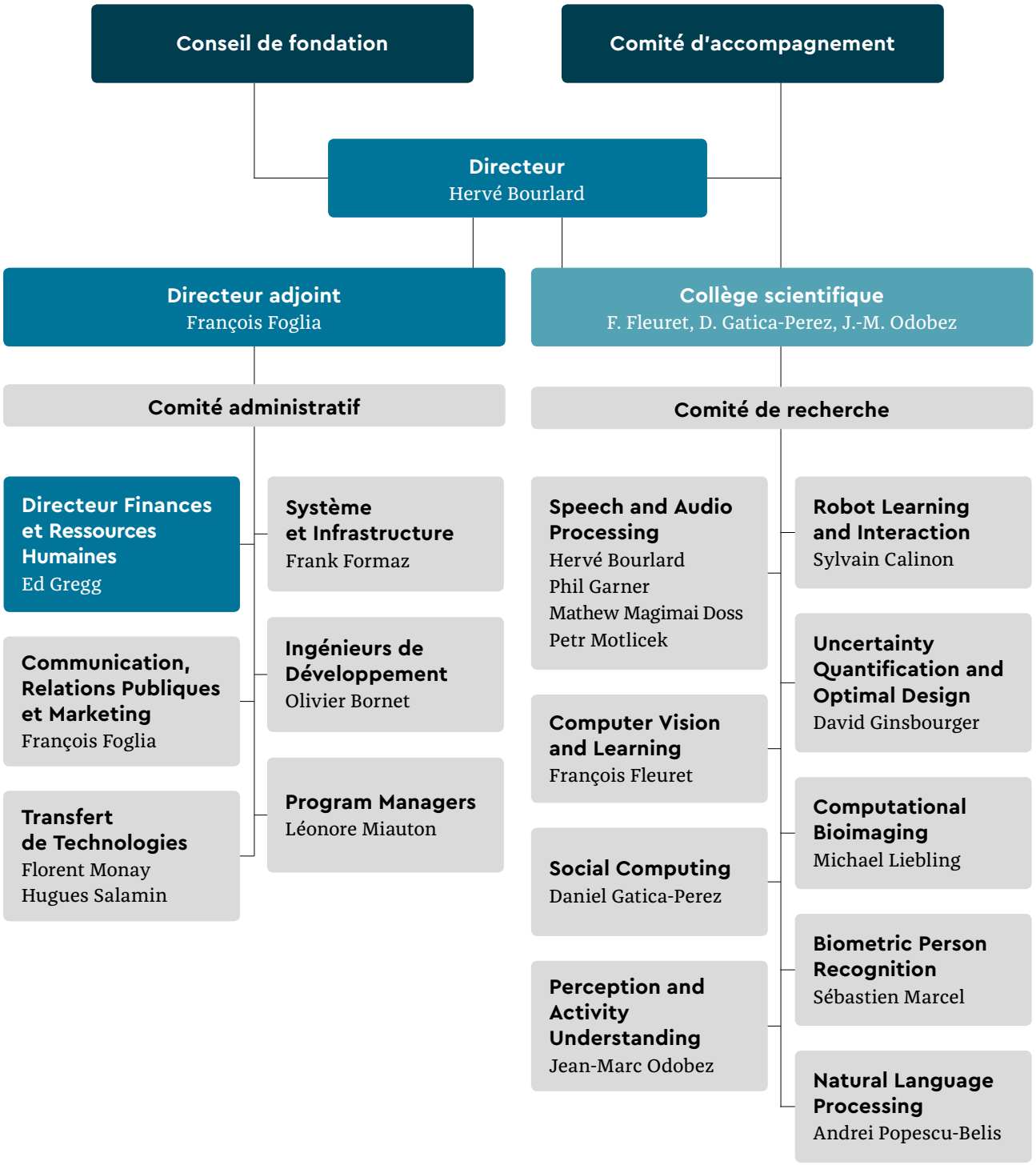
Divertissement et jeux

Jeux en réseau multilingues, jeux familiaux à distance (« togetherness »).

Domotique

Analyse de scènes multi-senseurs, analyse du comportement humain, adaptation au comportement humain (chauffage, etc.), sécurité ménagère.

Organigramme opérationnel



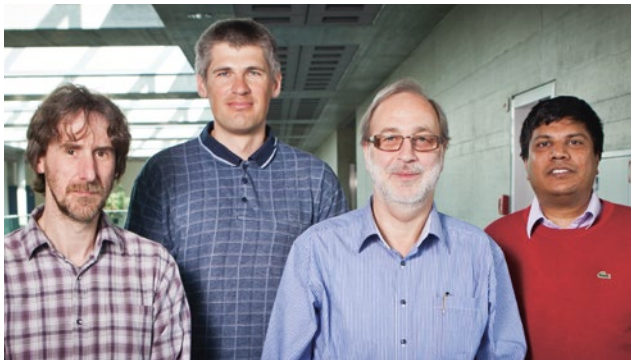
Groupes de recherche

À l'Idiap, les travaux sont menés au sein de neuf groupes de recherche. Celui concernant le 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).

Speech & Audio Processing

Prof. Hervé Bourlard (3e depuis la gauche),
Dr Phil Garner, Dr Petr Motlicek, Dr Mathew Magimai Doss
H-index respectifs*: PG: 22, PM: 15, HB: 57, MMD: 23

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).



Social Computing

Prof. Daniel Gatica-Perez
H-index: 50

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

Dr François Fleuret
H-index: 26

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 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.

Perception & Activity Understanding

Dr Jean-Marc Odobez

H-index: 37

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 multimédias constituent les principaux domaines d'application.

Uncertainty Quantification and Optimal Design

Dr David Ginsbourger

H-index: 18

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.

Biometric Person Recognition

Dr Sébastien Marcel

H-index: 38

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 (anti-spoofing) 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.

Robot Learning & Interaction

Dr Sylvain Calinon

H-index: 33

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.

Computational Bioimaging Group

Dr Michael Liebling

H-index: 19

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.

Natural Language Processing

Dr Andrei Popescu-Belis

H-index: 22

Le groupe Traitement automatique des langues étudie comment l'analyse des textes au niveau sémantique et pragmatique permet d'améliorer les performances dans deux tâches principales qui sont la traduction automatique et la recherche d'informations. Le groupe s'intéresse en particulier à la façon dont la recommandation d'informations depuis des bases de données multimédias en réseau peut être améliorée en utilisant la structure du réseau et le contenu informationnel de ses nœuds.

*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.

Conseil de fondation

Le Conseil de fondation assume la responsabilité de gestion économique et financière de l'institut de recherche. Il 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. Jean-Daniel Antille, Vice-Président
Responsable de l'Antenne Régions
Valais romand

Prof. Karl Aberer, Vice-Président
Vice-Président pour les systèmes
d'information, École polytechnique
fédérale de Lausanne

M. Marc-André Berclaz
Directeur opérationnel de l'Antenne
EPFL Valais-Wallis

M. Stefan Bumann
Chef de service des Hautes Ecoles (SHE)

M. Marc-Henri Favre
Président de la Ville de Martigny

M. Patrick Furrer
Vice-recteur Recherche
et Innovation, HES-SO

M. Jean-René Germanier
Encaveur et œnologue,
Cave Jean-René Germanier

M. Jordi Montserrat
Directeur régional de Venturelab

Prof. Christian Pellegrini
Doyen du Conseil de fondation de l'Idiap
Professeur honoraire à la Faculté des
sciences de l'Université de Genève

M. Dominique Perruchoud
Président du Conseil
d'administration de Cimark

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

Prof. Dr Vanessa Evers
Full Professor of Human Media Interaction, University of Twente, NL

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

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

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

Prof. Johanna Moore
University of Edinburgh, Head of the School of Informatics, Director of the Human Communication Research Center, UK

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

Collaborateurs

PERSONNEL SCIENTIFIQUE

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

Oya Aran, chercheuse associée, Turquie, 2009
Afsaneh Asaei, postdoctorante, Iran, 2008
Dario Azzimonti, doctorant, Italie, 2015
Sushil Bhattacharjee, chercheur associé, Suisse, Renens, 2016
Joan Isaac Biel Tres, postdoctorant, Espagne, 2008
Hervé Bourlard, directeur, Belgique/Suisse, Saxon, 1997
Sylvain Calinon, chercheur, Suisse, Martigny, 2014
Gulcan Can, doctorante, Turquie, 2013
Olivier Canévet, doctorant, France, 2012
Tatjana Chavdarova, doctorante, Macédoine, 2014
Ivana Chingovska, doctorante, Macédoine, 2015
Tiago De Freitas Pereira, doctorant, Brésil, 2014
Subhadeep Dey, doctorant, Inde, 2014
Pranay Dighe, doctorant, Inde, 2014
Marc Ferras Font, chercheur associé, Espagne, 2011
François Fleuret, chercheur senior, Suisse, Yverdon, 2007
Kenneth Alberto Funes Mora, postdoctorant, Costa Rica, 2011
Philip Garner, chercheur senior, Grande-Bretagne, 2007
Daniel Gatica-Perez, chercheur senior, Mexique, 2002
David Ginsbourger, chercheur senior, France, 2015
Ioannis Havoutis, postdoctorant, Grèce, 2015
Weipeng He, doctorant, Chine, 2016
Alexandre Heili, postdoctorant, France, 2015
Guillaume Heusch, chercheur associé, Suisse, Lausanne, 2015
Ivan Himawan, postdoctorant, Australie, 2016
Pierre-Edouard Honnet, doctorant, France, 2012
Rui Hu, postdoctorante, Chine, 2013
David Imseng, postdoctorant, Suisse, St. German, 2009
Christian Jaques, doctorant, Suisse, Lausanne, 2016
Noémie Jaquier, doctorante, Suisse, Vevey, 2016
Cijo Jose, doctorant, Inde, 2014
Alain Komaty, postdoctorant, France, 2016
Pavel Korshunov, postdoctorant, Estonie, 2015
Tipaluck Krityakierne, postdoctorant, Thaïlande, 2015
Ilja Kuzborskij, doctorant, Lettonie, 2012
Alexandros Lazaridis, chercheur associé, Grèce, 2012
Nam Le, doctorant, Vietnam, 2015
Rémi Lebret, doctorant, France, 2012
Joël Legrand, doctorant, France, 2012
Yann Lepoittevin, postdoctorant, France, 2016
Michael Liebling, chercheur senior, Suisse, St-Sulpice, 2015
Ngoc-Quang Luong, postdoctorant, Vietnam, 2014
Srikanth Madikeri, postdoctorant, Inde, 2013
Mathew Magimai Doss, chercheur, Inde, 2007
Parvaz Mahdabi, postdoctorant, Iran, 2014
Sébastien Marcel, chercheur senior, Suisse, Martigny, 2000
Olivia Mariani, doctorante, Suisse, Muraz, 2016
Angel Martinez Gonzalez, doctorant, Mexique, 2016
James Newling, doctorant, Angleterre, 2013
Lesly Miculicich, doctorante, Pérou, 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, Royaume-Uni, 2013
Laurent Nguyen, postdoctorant, Suisse, St-Légier, 2011
Olegs Nikisins, postdoctorant, Lituanie, 2016
Jean-Marc Odobez, chercheur senior, Suisse, Clarens, 2001
Pedro Henrique Oliveira Pinheiro, doctorant, Brésil, 2012
Dimitri Palaz, doctorant, Suisse, Martigny, 2011
Nikolaos Pappas, doctorant, Grèce, 2012
Novi Patricia, doctorant, Inde, 2012
Xiaojiang Peng, postdoctorant, Chine, 2016
Trung Phan, doctorant, Vietnam, 2015
Emmanuel Pignat, doctorant, Suisse, Muraz, 2015
Andrei Popescu-Belis, chercheur senior, Suisse, Lausanne, 2007
Xiao Pu, doctorante, Chine, 2014
André Rabello Dos Anjos, chercheur associé, Brésil, 2010
Dhananjay Ram, doctorant, Inde, 2014
Ramya Rasipuram, postdoctorante, Inde, 2010
Marzieh Razavi, doctorante, Iran, 2013
Dairazalia Sanchez-Cortes, postdoctorante, Mexique, 2012
Darshan Santani, postdoctorante, Inde, 2012
Adrian Shajkofci, doctorant, Suisse, Martigny, 2016
Ajay Srinivasamurthy, postdoctorant, Inde, 2016
Ajay Tanwani, doctorant, Pakistan, 2015
Sibo Tong, doctorant, Chine, 2016
Sandrine Tornay, doctorante, Suisse, Vernayaz, 2016
Raphael Ullmann, doctorant, Suisse, Lausanne, 2012
Michael Villamizar, postdoctorant, Espagne, 2016
Jonas Steffen Walheim, doctorant, Allemagne, 2016
Yang Wang, postdoctorant, Chine, 2016
Yu Yu, doctorant, Chine, 2015

- **Ingénieurs de développement**

Philip Abbet, ing. de développement senior, Suisse, Conthey, 2006
Olivier Bornet, responsable ing. de développement, Suisse, Pont de la Morge, 2004
Milos Cernak, ing. de développement senior, Slovaquie, 2011
Samuel Gaist, ing. de développement, Suisse, Ardon, 2013
Salim Kayal, ing. de développement, Suisse, Vevey, 2011
Vasil Khalidov, ing. de développement senior, Russie, 2014
Christine Marcel, ing. de développement, Suisse, Martigny, 2007
Florent Monay, officier de transfert de technologie, Suisse, Monthey, 2008
Alexandre Nanchen, ing. de développement senior, Suisse, Martigny, 2008
Hugues Salamin, officier de transfert de technologie, Suisse, Dorénaz, 2014
Flavio Tarsetti, ing. de développement senior, Suisse, Martigny, 2008

PERSONNEL ADMINISTRATIF

Elisa Bovio, junior program manager et aide-comptable, Suisse, Martigny, 2015
Martina Fellay, program manager, Autriche, 2012
François Foglia, directeur adjoint, Suisse, Saxon, 2006
Edward Gregg, responsable financier, Etats-Unis, 2004
Léonore Miauton, responsable des program managers, Suisse, Chexbres, 2012
Sylvie Millius, secrétaire, Suisse, Vétroz, 1996
Nadine Rousseau, secrétaire, Suisse, Saxon, 1998

- **Ingénieurs système**

Bastien Crettol, administrateur système, Suisse, Martigny, 2005
Norbert Crettol, administrateur système senior, Suisse, Martigny, 2002
Cédric Dufour, ingénieur système, Suisse, Aigle, 2007
Frank Formaz, responsable des ingénieurs système, Suisse, Fully, 1998
Louis-Marie Plumel, administrateur système senior, France, 2011
Vincent Spano, webmaster, Suisse, Martigny-Combe, 2004

STAGIAIRES

Prénom, nom, origine, institution d'origine

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.

Samaneh Abbasi, Iran, EPF Lausanne
Christian Abbet, Suisse, EPF Lausanne
Krysten Ansermoz, Suisse, EPF Lausanne
Afroze Baqapuri, Pakistan, EPF Lausanne
Maxime Bourlard, Belgique, Université de Neuchâtel
Agata Cernakova, Slovaquie,
Kevin Chan, Etats-Unis, University of California (US)
LongTao Chen, Chine, Nanjing Univ. of Science and Technology (CN)
Fabien Crepon, Suisse, EPF Lausanne
Jaden Diefenbaugh, Etats-Unis, Oregon State University (US)
Teodors Eglitis, Lituanie, Riga Technical University (LT)
Christopher Finelli, Suisse, EPF Lausanne
Nikol Guljelmovic, Pays-Bas, Delft univ. of Technology (NL)
Maryam Habibi, Iran, Institut de recherche Idiap
Gil Luyet, Suisse, Université de Fribourg
Alan Maitre, Suisse, EPF Lausanne
Zohreh Mostaani, Iran, Université Ozyegin (TR)
Terry Niederhauser, Suisse, HES-SO
Amrutha Prasad, Inde, Visvesvaraya Technological University (ID)
Nilavan Sathasivam, Suisse
Amrollah Seifoddini Banadkooki, Iran, ETHZ, Zurich
Dominique Tao, Suisse, EPF Lausanne
Véronique Thurre, Suisse
Kelly Tiraboschi, Italie, Université de Lausanne
Bogdan Vlasenko, Ukraine, Otto-von-Guericke Universität (GER)
Jing Yang, Chine, ETH Zurich

VISITEURS

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

Chercheurs ou industriels, les visiteurs ne passent que quelques jours ou quelques semaines à l'institut, les uns pour renforcer les liens interinstitutionnels, les autres pour prendre la mesure des travaux qui s'effectuent dans l'institut.

Yuxiang Li, Chine, EPFL
Catharine Oertel, Allemagne, Université de Bielefeld (DE)

finances

18 — 21

Compte d'exploitation (CHF)

PRODUITS	2015	2016	%
Commune de Martigny	700 000	**700 000**	6.5%
État du Valais	1 720 000	**1 720 000**	16.1%
Confédération	2 467 300	**2 423 700**	22.6%
Loterie Romande	100 000	**87 000**	0.8%
Projets Fonds national suisse de la recherche scientifique	1 715 249	**1 537 844**	14.3%
Projets Fondation Hasler	606 689	**175 781**	1.6%
Projets européens	1 395 595	**1 404 423**	13.1%
Projets The Ark	225 070	**245 380**	2.3%
Projets CTI	388 170	**638 467**	6%
Contribution EPFL	72 000	**72 000**	0.7%
Financement industriel	538 818	**414 669**	3.9%
Autres financements / Produits exceptionnels	396 274	**738 592**	6.9%
Provisions	—	**555 000**	5.2%
TOTAL DES PRODUITS	**10 325 165**	**10 712 856**	**100.00 %**

CHARGES			
Frais de personnel	7 884 132	**7 948 334**	74.2%
Formation et déplacements	371 774	**427 372**	4%
Partenaires externes	167 114	**102 292**	1%
Informatique : matériel et maintenance	185 680	**253 410**	2.4%
Frais administratifs	148 148	**166 113**	1.6%
Promotion et communication	45 468	**45 769**	0.4%
Loyer et charges	791 284	**709 229**	6.6%
Amortissement	263 690	**363 090**	3.4%
Charges exceptionnelles	—	**—**	—
Provisions	455 000	**685 000**	6.4%
TOTAL DES CHARGES	**10 312 290**	**10 700 609**	**100%**
RÉSULTAT D'EXPLOITATION	**12 875**	**12 247**	**—**

Commentaires sur les comptes 2016

Subventions
Confédération, Canton, Commune
(En milliers de francs suisses)

ANNÉES	2014	2015	2016	2017*
Confédération	2550	2467	2424	2418
Canton	1720	1720	1720	1645
Commune	700	700	700	700

*Budget

Pour la cinquième année consécutive, l'Idiap termine l'exercice 2016 avec des revenus supérieurs à 10 millions.

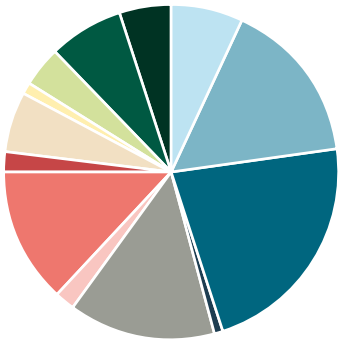
La part des subventions publiques est inférieure à 50% de l'ensemble des revenus et respecte, ainsi, l'objectif stratégique fixé par la Direction. Depuis quelques années, l'Idiap met l'accent sur le développement de ses relations industrielles cantonales, nationales et internationales et le transfert de technologie. Force est de constater que ces efforts portent leurs fruits. Plus de 10% du budget est consacré à des projets en lien avec l'économie (sources de financement : TheArk, CTI, industriel).

Les frais de personnel (salaires + charges sociales) forment environ ¾ de la totalité des charges.

L'Idiap boucle ses comptes 2016 sur un excédent de 12'247 francs.

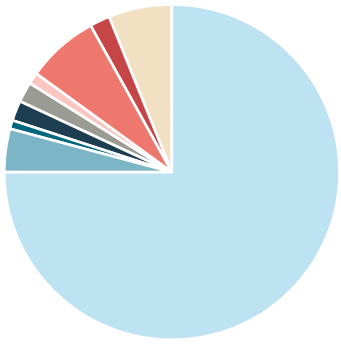
Répartition des sources de financement

- 7% Commune de Martigny
- 16% Etat du Valais
- 22% Confédération
- 1% Loterie Romande
- 14% Projets Fonds national suisse de la recherche scientifique
- 2% Projets Fondation Hasler
- 13% Projets européens
- 2% Projets The Ark
- 6% Projets CTI
- 1% Contribution EPFL
- 4% Financement industriel
- 7% Autres financements, Produits exceptionnels
- 5% Provisions



Répartition des charges

- 75% Frais de personnel
- 4% Formation et déplacements
- 1% Partenaires externes
- 2% Informatique : matériel et maintenance
- 2% Frais administratifs
- 1% Promotion et communication
- 7% Loyer et charges
- 2% Amortissement
- 6% Provisions



Bilan (CHF)

ACTIFS	31.12.2015	31.12.2016
Liquidités	3 467 573.13	5 313 117.83
Débiteurs	319 121.25	130 394.85
Actifs de régularisation et divers	622 013.91	434 238.84
TOTAL ACTIFS CIRCULANTS	**4 408 708.29**	**5 877 751.52**
Mobilier et matériel informatique	434 031.64	590 125.14
Participations financières	10 000.00	10 000.00
TOTAL ACTIFS IMMOBILISÉS	**444 031.64**	**600 125.14**
TOTAL ACTIFS	**4 852 739.93**	**6 477 876.66**

PASSIFS	31.12.2015	31.12.2016
Créanciers	349 413.85	213 106.26
Passifs de régularisation	1 878 638.93	3 747 835.83
Provisions	1 305 000.00	1 085 000.00
TOTAL FONDS ÉTRANGERS	**3 533 052.78**	**5 045 942.09**
Capital	40 000.00	40 000.00
Réserve spéciale	1 100 000.00	1 200 000.00
Résultat reporté	166 812.11	179 687.15
Résultat net	12 875.04	12 247.42
TOTAL FONDS PROPRES	**1 319 687.15**	**1 431 934.57**
TOTAL PASSIFS	**4 852 739.93**	**6 477 876.66**

recherche

22 — 29

L'enseignement à distance 3.0

Les parcours professionnels, aujourd'hui de moins en moins linéaires, ainsi que les nouvelles technologies imposent une mise à jour de plus en plus rapide des connaissances pour qui souhaite maintenir ses compétences et sa compétitivité. L'enseignement continu à distance offre une alternative intéressante à qui souhaite poursuivre une formation tout en la conciliant avec ses vies familiale et professionnelle.

L'enseignement à distance permet de suivre un enseignement de qualité et reconnu, tout en étant adapté aux besoins de chacun et sans avoir à ajouter à un agenda déjà bien rempli les déplacements nécessaires à retourner sur les bancs d'école. Ce type d'enseignement continu de qualité ne peut se faire qu'au moyen d'une plate-forme robuste et conviviale, favorisant au maximum les interactions entre les étudiants ainsi qu'entre ces derniers et leurs professeurs.

Active depuis plus de 20 ans dans ce domaine, et comptant actuellement 1500 étudiants actifs, UniDistance s'est alliée avec l'Idiap et sa spin-off Klewel dans le cadre du projet eLearning-Valais 3.0. Ce projet vise à mettre au point une solution inédite en matière de production et d'utilisation de ressources multimédia et interactives pour la formation, tirant profit de la complémentarité des compétences des trois partenaires :

- UniDistance: pédagogie universitaire,
- Klewel : acquisition, indexation et partage d'exposés multimédia,
- Idiap : analyse des données multimédia.

Allant au-delà des « MOOCs » (Massive Open Online Courses), le projet eLearning-Valais 3.0 vise à intégrer au sein d'une solution numérique unique un ensemble de fonctionnalités existant déjà séparément :

- production simplifiée d'enregistrements,
- navigation intelligente,
- indexation multilingue des enregistrements,
- génération de mots-clés pour retrouver facilement un document ou une section particulière du curriculum,
- partage des annotations de cours avec ses pairs pour aller rapidement à l'essentiel,
- suggestion de documents pertinents.

A long terme, ce véritable couteau suisse pour l'enseignement et les apprentissages à l'ère numérique permettra aussi de mieux comprendre les processus de collaboration dans les apprentissages à distance. Au niveau régional, le projet eLearning-Valais 3.0 permettra ainsi au canton du Valais de renforcer sa position de leader dans le domaine de l'enseignement continu à distance, tout en répondant aux attentes sociales et aux besoins des étudiants et différents milieux professionnels.

Klewel
the webcasting company

FORMATION UNIVERSITAIRE
A DISTANCE, SUISSE

L'Idiap maintient le cap !

Comme pour bien d'autres acteurs de la politique fédérale, le monde de la recherche helvétique vit sur un grand rythme de quatre ans pour lesquels les budgets sont discutés à l'avance. Si le prochain exercice concerne la période 2017–2020, c'est bien évidemment entre 2015 et 2016 que se sont déroulées des négociations importantes concernant l'Idiap.

Pour la prochaine période quadriennale, le budget alloué par la Confédération, au travers de l'article 15(b) de la loi fédérale sur l'encouragement de la recherche et de l'innovation (art. 15, al. 3, let. b, LERI) ne subira pas de changement. Pour l'Idiap, il se montera à nouveau à près de 10mio CHF sur 4 ans (2017-2020).

Bien que le budget alloué à la recherche en général, et aux institutions relevant de l'article 15 en particulier soit en très légère hausse, la Confédération a décidé de donner la priorité au renforcement des centres de compétences technologiques (art. 15, al. 3, let. c, LERI), à la création de nouveaux centres, ainsi qu'à la consolidation des infrastructures de recherche.

Alors que bon nombre d'institutions doivent faire face à une baisse de leurs crédits, l'Idiap peut donc se targuer d'un statu quo financier. Pour Hervé Bourlard, directeur de l'Idiap : « Dans un contexte plutôt tendu pour les finances publiques, je vois dans le maintien de notre enveloppe budgétaire fédérale un renouvellement de la confiance envers notre institut. Et cette enveloppe devrait être suffisante sur le moyen terme pour nous permettre de soutenir de façon significative le développement de l'institut. »

« Une confiance renouvellée. »

Poursuite de la diversification

Ainsi, et avec le soutien indéfectible de l'Etat du Valais et de la Ville de Martigny, l'Idiap peut poursuivre sa diversification et mettre au concours, comme prévu, deux nouveaux postes de chercheurs permanents, dont un dans le domaine de « Energy Informatics ». Ceci devrait permettre un rapprochement avec le Centre de Recherche en Energie Municipale (CREM) de Martigny, et de participer au développement de l'écosystème valaisan dans le domaine de l'énergie en collaboration avec des partenaires forts comme l'EPFL-Valais-Wallis et la HES-SO.

L'Idiap et l'EPFL

Cette nouvelle période budgétaire va de pair avec la signature, en mars 2017, du renouvellement du Plan de développement conjoint avec l'EPFL. Celui-ci vise à répondre au souhait de la Confédération de voir l'Idiap et d'autres instituts de recherche affiliés à l'EPFL intégrer le réseau académique suisse. Ce nouveau plan se veut beaucoup plus ambitieux, engageant les deux partenaires à collaborer pour le développement national de certaines de leurs activités communes.

La biométrie au service de la cybersécurité

Alors que nous vivons dans un environnement de plus en plus connecté, les questions de cybersécurité n'intéressent pas uniquement les spécialistes. Afin de s'y retrouver dans la gestion des mots de passe de plus en plus nombreux la plupart d'entre nous font généralement au plus simple : un mot de passe peu complexe, pour plusieurs usages, parfois même écrit quelque part pour être sûr de ne pas l'oublier… Du pain bénit pour les cybercriminels !

A cet égard, la biométrie offre des solutions simples. En effet, l'être humain possède de nombreuses particularités physiques qui le rendent unique : empreintes digitales, iris, forme du visage, réseau veineux, voix… C'est précisément dans ce domaine que le Centre de Recherche et d'Evaluation en Sécurité Biométrique de l'Idiap développe ses activités. Sous la houlette du Dr Sébastien Marcel, ce centre regroupe actuellement quatre doctorants, trois postdoctorants et trois collaborateurs scientifiques.

« Nous développons, en particulier, des technologies d'authentification et d'identification, explique Sébastien Marcel. Dans le premier cas, une signature physique est comparée à une référence, ce qui permet de vérifier si oui ou non on est bien celui que l'on prétend être. Dans le deuxième cas, il s'agit de trouver la bonne corrélation entre une signature anonyme et une autre, connue et contenue dans une base de donnée. »

Si le traitement de l'information est développé spécifiquement à l'Idiap, des collaborations avec d'autres, écoles, instituts ou entreprises sont indispensables pour développer et intégrer les capteurs appropriés (par ex. lecteurs d'empreintes, de veines).

Un autre aspect de la cybersécurité a trait aux attaques visant à usurper une identité. Pour ce qui est de la biométrie, on peut facilement imaginer se faire passer pour quelqu'un d'autre en utilisant une photo ou un enregistrement de la voix d'une personne cible fait à son insu ou trouvé sur internet. Il pourrait également provenir du piratage d'une banque de données biométriques. Il s'agit donc là de développer des technologies permettant de distinguer une signature (par exemple : vocale) originale d'une copie.

En 2016 l'Idiap et son Centre de compétence en Biométrie a organisé deux événements très différents. En janvier, l'institut a accueilli le groupe de travail « Biometrics », une première pour la Suisse et le Valais, et a vu ainsi se réunir 87 spécialistes du monde entier venus débattre des normes et de la mise en œuvre des standards ISO autour de la thématique de la biométrie. Dans un autre contexte, en octobre, un stand organisé à la Foire du Valais, qui a cette année attiré plus de 200 000 visiteurs, a permis de valoriser les activités en biométrie de l'institut et, également, d'enregistrer les données biométriques de plus de 500 volontaires ; données qui seront utilisées dans le cadre de futurs projets de recherche.

Baxter, le robot qui vous habille

Aider une personne entravée dans sa mobilité à s'habiller ou à se chausser, tel est l'objectif que vise le groupe de recherche du Dr Sylvain Calinon avec le projet i-Dress. Celui-ci se base sur le robot Baxter qui sert également à tester les algorithmes de projets aussi divers qu'un robot sous-marin ou une main artificielle.

Si un sportif blessé peut logiquement espérer voir sa mobilité s'améliorer rapidement, il en va autrement pour une personne âgée, chez qui elle peut même se péjorer inexorablement. Un robot d'aide à l'habillement devrait donc être idéalement capable de s'adapter à l'évolution physique de son patient, de manière à lui apporter en tout temps une aide aussi personnalisée que possible. De plus, pour qu'il soit utilisable dans la vie quotidienne, il doit se montrer relativement souple, par exemple en suivant les déplacements de son patient, sans le retenir ou le forcer. Enfin, le personnel soignant doit, quant à lui, être en mesure de maîtriser la partie « médicale » de la programmation, qui se doit d'être la plus intuitive possible.

Au vu de ces nombreux défis, la plate-forme robotique Baxter s'avère le candidat idéal. Il s'agit d'un robot d'une grande souplesse, auquel il est possible d'enseigner un mouvement par imitation, c'est-à-dire en le guidant physiquement, sans avoir recours à un langage de programmation.

Au bénéfice d'un financement CHIST-ERA, le projet i-Dress a débuté en 2016 et réunit trois équipes de recherche. Le groupe de l'Idiap, dirigé par Sylvain Calinon, s'occupe des algorithmes ayant trait à l'apprentissage par imitation. Des chercheurs de l'Université polytechnique de Catalogne travaillent à déterminer la position des vêtements ou chaussures à déplacer. Ces objets, relativement grands, mais surtout très déformables et donc jamais tout à fait identiques, s'avèrent relativement compliqués à percevoir. Enfin, une équipe de l'Université de « West of England », à Bristol, travaille sur le versant humain du projet, à savoir comment les futurs patients souhaitent interagir avec le robot.

Conçu il y a quatre ans, à la base comme robot industriel, Baxter est utilisé dans la recherche, où il représente un outil de prototypage et de test d'algorithmes idéal. Dans cette configuration, Sylvain Calinon l'utilise également dans le cadre du projet de robot sous-marin télé-opéré DexROV ainsi que pour des recherches sur une main artificielle (TACT-HAND).

L'Idiap outre-Atlantique

De prestigieuses collaborations avec les USA

Au niveau du financement de ses projets de recherche, l'Idiap ne se limite pas à la Suisse et à l'Europe. Si l'institut bénéficie du soutien du Fonds national suisse, de la Commission pour la technologie et l'innovation et des programmes de recherche H2020 de l'Union européenne, il est régulièrement un partenaire de choix dans le cadre de programmes de recherche des USA.

En 2016, les spécialistes de la biométrie du groupe de Sébastien Marcel se sont vus intégrés dans deux projets financés par le « Defense Advanced Research Projects Agency » (DARPA) et l' « Intelligence Advanced Research Projects Activity » (IARPA). Ces deux organismes, liés au Département américain de la défense et du renseignement, ont un budget annuel de recherche de l'ordre de six milliards de dollars à eux deux. Leur but est de financer des projets à haut risque mais également à fort potentiel, très spécifiques, susceptibles de déboucher sur des avancées technologiques de premier ordre. On leur doit, entre autres, l'Internet, la miniaturisation des systèmes GPS et l'assistant vocal Siri. Ces technologies ont toutes en commun d'avoir d'abord été développées pour un usage militaire avant d'être ouvertes au grand public.

Entre prestige et risques

Les appels à candidature de ces deux institutions ont en commun d'être très sélectifs et très concurrentiels, si bien que le nombre de laboratoires susceptibles d'être retenus est relativement faible. L'Idiap et son Centre de recherche et d'évaluation en sécurité biométrique ont eu la chance de se voir invités à rejoindre ce club très fermé, un excellent signe au niveau de leur réputation internationale et de la réussite du pari pris par l'Idiap de développer ses activités en biométrie.

Si les montants alloués à ces projets de grande envergure sont conséquents - de l'ordre de 20-30 millions -, leur financement intégral n'est jamais garanti. En effet, chaque projet met en concurrence plusieurs équipes travaillant sur la même thématique et il arrive régulièrement que les moins avancées se voient éliminées de la partie, une philosophie très différente des organismes de financement de la recherche suisses et européens. Une autre inconnue, d'actualité celle-ci, tient aux changements d'administration, car la nouvelle équipe dirigeante peut facilement décider de revoir l'agenda de la précédente et de réorienter les axes de sa stratégie. Pour faire face, entre autres, à ces aléas, la direction de l'Idiap tient à jour une analyse de risque afin d'éviter que l'arrêt d'un projet mette financièrement l'institut en danger.

Cybersécurité et biométrie à l'honneur

Les travaux pour le projet DARPA ont déjà débuté. Nommé SAVI, pour « Spotting Audio-Visual Inconsistencies » et financé dans le cadre du programme MediFor « Media Forensics », ce projet vise à détecter les manipulations, faites à des fins de propagande, des sources audiovisuelles, et à lutter contre la désinformation (« fake news »). Concrètement, il s'agit de parvenir à déterminer s'il est possible de détecter des différences entre ce que l'on voit dans un document audiovisuel et ce qu'on y entend. S'il ne s'agit pas de biométrie à proprement parler, les technologies employées font partie du thème de la cybersécurité. Ce projet est dirigé par SRI International et réalisé en partenariat avec l'Université d'Amsterdam.

A la fin de l'année 2016, l'équipe de Sébastien Marcel a également reçu la confirmation de la participation de l'Idiap et de son Centre en biométrie au programme Odin de l'IARPA, dont le but est de lutter contre les attaques de présentation, ou « spoofing », visant à tromper les systèmes de sécurité biométrique. Approché par trois équipes concurrentes pour y participer, l'Idiap a fait le choix de l'équipe gagnante et collaborera sur un projet du programme Odin qui permettra de recruter deux nouveaux chercheurs.

Conférence internationale en « deep learning »

Du 4 au 6 juillet 2016, l'Idiap a accueilli 300 chercheurs et étudiants à l'occasion d'un workshop sur les méthodes et outils du deep learning. Une discipline en plein essor, appelée à modifier fortement notre quotidien.

En octobre 2015, l'ordinateur AlphaGo bat le champion européen de go sur le score sans appel de 5:0. Quelques mois plus tard, il s'impose à nouveau, 4:1, face au champion du monde. Une prouesse que nombre de spécialistes jugeait jusque-là impossible. En effet, le jeu de go est infiniment plus complexe que les échecs : quand Deep Blue battait Kasparov, en 1997, l'informatique n'avait raison que de médiocres joueurs de go.

Derrière cet exploit se cache ce que les spécialistes appellent le deep learning. Une discipline née dans les années 1950 sous le nom de réseaux neuronaux et qui, grâce au développement de la puissance de calcul des ordinateurs, connaît actuellement un véritable renouveau. L'Idiap, depuis longtemps actif dans ce domaine, est à l'origine de Torch, un des principaux outils de développement pour cette technologie encore utilisé et optimisé chez Facebook.

Au retour d'une des plus importantes conférences dédiées au deep learning, François Fleuret, responsable du groupe de Computer Vision and Learning de l'Idiap, décide d'organiser un workshop pour ses étudiants. Grâce à ses connexions, il y invite des spécialistes du software (Facebook et Google) comme du hardware (NVIDIA et AMD). En complément à ce programme, les conférences du Prof. Yoshua Bengio, de l'Université de Montréal (3e à partir de la gauche sur la photo), ont pour thème les dimensions historiques et contextuelles du deep learning. L'enthousiasme pour ce workshop dépasse vite le cercle des étudiants et c'est finalement un public international de 300 personnes qui se retrouve à Martigny.

« De 1990 à 2010 environ, les réseaux neuronaux avaient perdu de leur popularité, explique François Fleuret. La recherche en intelligence artificielle portait quasi exclusivement sur des techniques alternatives, plus accessibles à une analyse mathématique. Mais ces réseaux permettent d'exploiter efficacement les ordinateurs modernes, en particulier les unités de calcul graphiques qui vont des centaines de fois plus vite que les processeurs classiques. Ces technologies permettent d'entraîner les ordinateurs avec les immenses quantités de données digitales disponibles grâce à internet et nous emmènent aujourd'hui aux portes d'une révolution technologique qui va bouleverser notre quotidien. »

Aujourd'hui, la reconnaissance automatique est un des grands thèmes du deep learning. Des données abstraites d'un fichier image, il s'agit d'extraire du sens. « Nourrie de données, la machine apprend petit à petit à différencier une vache d'un cheval, poursuit François Fleuret. Un processus analogue à celui qui a permis à AlphaGo de devenir un maître : abreuvé de milliers de parties, il en a assimilé les stratégies, puis s'est entraîné en jouant contre lui-même, avant de battre le meilleur d'entre nous. »

Moi, Pepper, robot humanoïde

Les chercheurs de l'Idiap ont fait l'acquisition de la plate-forme robotique Pepper. Leur but ? En faire un robot social capable d'interagir avec les humains dans un environnement complexe.

On ne présente plus Nao aux familiers de l'Idiap et de la robotique. Depuis peu, ce petit robot humanoïde, autonome et programmable, développé par Aldebaran Robotics, société reprise en 2015 par Softbank Robotics, a un grand frère, Pepper, qui mesure 140 cm, contre 48 pour Nao. Autre différence anatomique : si Nao se déplaçait sur ses deux jambes, les créateurs de Pepper ont renoncé à la bipédie au profit de trois roues lui conférant une plus grande stabilité et une meilleure mobilité. Enfin, une tablette disposée sur son tronc lui permet d'interagir plus facilement avec ses vis-à-vis. Tant au niveau de son apparence que de ses fonctionnalités, il est très différent d'un robot comme Baxter (voir article p. 26). Si ce dernier a des origines industrielles, Pepper est lui un robot social appelé à interagir avec les humains.

Robotique sociale

C'est précisément dans cette optique que Pepper a intégré, en été 2016, le groupe de recherche « Perception and Activity Understanding » du Dr Jean-Marc Odobez. Il y servira de plateforme d'expérimentation pour le nouveau projet européen MuMMER (MultiModal Mall Entertainment Robot), mené par un consortium de sept institutions. Le but de ce projet est de développer un robot capable d'interactions sociales avec les clients d'un grand centre commercial. A la fois divertissement, homme-sandwich, ou service à la clientèle, Pepper devra y opérer dans un environnement complexe au moyen de ses sens, ou plutôt de ses capteurs.
Dans ce projet, le groupe de Jean-Marc Odobez est justement responsable de l'aspect « perception des personnes » : localisation, prise de parole et communication non verbale (signaux sociaux). A l'heure actuelle, les chercheurs travaillent à analyser les capacités (et les limites) des capteurs de cette nouvelle plate-forme et à en améliorer la faculté à saisir des émotions.

réseau

30 — 37

Le retour au Pays

Favoriser le retour des Valaisan-ne-s dans leur canton d'origine après l'achèvement de leurs études, tel est le but du projet « Valais-Wallis Ambition ». Une façon aussi de faire bénéficier le tissu économique cantonal des investissements consentis par l'Etat en matière de recherche et formation.

Ne disposant pas de sa propre université, le canton du Valais se voit confronté à une problématique de fuite des cerveaux quand bien même il participe financièrement à la formation de plusieurs milliers d'étudiants par année. En effet, selon un rapport publié en novembre 2016 par le Service des hautes écoles du Département de la formation et de la sécurité, l'Etat valaisan a consacré, en 2015, 61,5 millions de francs pour les 5712 étudiants inscrits dans une université, une EPF ou une HES hors du canton (et 55 millions pour les 5599 étudiants suivant une formation en Valais). Se pose également la question du retour de ces étudiants, et particulièrement pour ceux d'entre eux qui souhaiteraient poursuivre leur parcours académique dans une haute école ou un institut de recherche universitaire en Valais.

Afin de favoriser ce retour, l'Idiap a lancé en 2016 l'initiative « Valais-Wallis Ambition ». Celle-ci vise à offrir aux jeunes ayant effectué leur scolarité obligatoire et secondaire en Valais une opportunité de compléter leur formation universitaire en réalisant un doctorat ou un postdoctorat au sein de l'Institut octodurien. La sélection s'opère sur dossier, en fonction de la qualité scientifique des candidats et de leur adéquation aux domaines de recherche de l'Idiap. Pour François Foglia, directeur adjoint de l'Idiap, « le but de cette initiative n'est pas de financer des projets de recherche sur le long terme, mais de mettre en lien nos projets de recherche en cours ou à venir et des candidats désireux de rentrer au Pays. Il est également possible d'imaginer soumettre un nouveau projet de recherche autour des compétences d'un candidat accepté à l'Idiap dans le cadre de ce programme. Valais-Wallis Ambition sert alors à financer, sur le moyen-terme (période tampon), le salaire du (post)doctorant en attendant le démarrage officiel d'un projet soutenu par des fonds compétitifs, FNS, Fondation Hasler, CTI ou européens, par exemple ».

« Opportunité pour les jeunes valaisan–ne-s. »

Ce programme a immédiatement rencontré un vif succès. En 2016, trois doctorants ont déjà été engagés sous l'égide de « Valais-Wallis Ambition ». Leur portrait se trouve dans les pages « Visages » de ce rapport.

L'Idiap fête ses 25 ans !

Pour célébrer son premier quart de siècle, l'Idiap est allé à la rencontre de ses partenaires et de la population. Du 1e au 3 septembre, trois manifestations distinctes lui ont ainsi permis de présenter ses activités et son savoir-faire à près de 700 personnes.

Les festivités ont débuté le jeudi 1er septembre par la tenue d'une Journée Innovation. Principalement destinée aux acteurs académiques et industriels, elle a réuni près de 200 personnes auxquelles l'ensemble des groupes de recherche de l'Idiap ont présenté leurs différents projets.

Ensuite, durant plus de trois heures, les invités se sont promenés dans les couloirs de l'institut, où étaient installés plus de 30 démonstrations ainsi que des posters illustrant certains tournants historiques de l'institut. Ils ont pu expérimenter au plus près toute la palette des activités scientifiques et de transfert de technologie de l'Idiap.

La journée s'est achevée par une conférence de Silvino Schlickmann, responsable des activités de recherche sur la cybercriminalité auprès d'INTERPOL, sur le sujet des menaces émergentes bien réelles issues du monde virtuel.

L'Idiap a profité de la mise en place de toute cette organisation pour inviter, le vendredi 2 septembre, les huit membres de son nouveau Conseil Scientifique à l'occasion de leur rencontre bisannuelle. Ces experts internationaux ont ainsi eux aussi pu s'immerger au cœur des travaux de recherche, avant leur scéance plus technique.

Enfin, le samedi 3 septembre, c'est la population – principalement valaisanne - qui était invitée à franchir les portes grandes ouvertes de l'Idiap. Ce ne sont pas moins de 450 personnes, dont de nombreux enfants et familles, qui ont répondu présent et se sont plongés dans l'univers de l'intelligence artificielle, de la robotique, du traitement du language ou encore de la biométrie.

En complément à ces manifestations exceptionnelles, l'institut a également inauguré son « 5 à 6 », tenu le deuxième mercredi de chaque mois, où le public est invité à une présentation et discussion ouverte autour des différentes activités de recherche de l'Idiap. Compte tenu du succès de ces rencontres, elles continueront à être programmées en 2017.

idiap
Bienvenue!
Bienvenue!
idiap
idiap
idiap

idiap

De l'idée à la start-up

Lancé par l'Idiap en 2012, l'International Create Challenge (ICC) a pour but de permettre à de jeunes entrepreneurs de transformer leur idée en prototype commercial au terme de trois semaines d'immersion à l'institut. En 2016, pour sa 5e édition, ce programme unique a une nouvelle fois atteint sa cible. Quatre projets innovants se sont vus au final distingués.

Du 31 août au 21 septembre 2016, l'Idiap a accueilli huit équipes d'entrepreneurs en herbe - scientifiques, designers, et responsables marketing et communication - en provenance notamment de Suisse, de France, d'Espagne et du Brésil. A l'issue de l'ICC 2016, le jury, composé de personnalités du monde de la recherche, du capital-risque et de l'innovation, a récompensé quatre projets alliant innovation et haute qualité scientifique que sont Viprodev, Veintree, MeasureMe et Dental Simulator.

Participants ICC'2016

1. VIPRODEV

ÉQUIPE

- Dr Kadir Akin : postdoctorant au Microelectronic Systems Laboratory (LSM) de l'EPFL (Suisse)
- Georgel Vilsan : Master en Public Sector Management (Roumanie)
- Dr Pablo Garcia : PhD en Computer Systems Engineering de l'Université Complutense de Madrid (Espagne)
- Luciano Cerqueira Torres : PhD en Strategy, Program and Project Management de l'École Supérieure de Commerce de Lille (France)

OBJECTIF

VIPRODEV vise à développer un appareil stéréoscopique de mesure de distance à haute performance ainsi qu'une interface graphique pour ses utilisateurs. Nommé Depthor, il utilise une combinaison de caméras RGB et de profondeur pour des applications de traitement d'images.

Equipe Viprodev

2. VEINTREE

ÉQUIPE

- Christophe Bron : Dr med. MSc. actuellement gérant de VTIP Sàrl, et président du Groupe Veintree (France)
- Mathieu Rossi : ingénieur informaticien et gérant de R4M (France)
- Didier Launay : vice-président de XMP Entrepreneur (France)
- Luc Rodrigues de Magalhaes : développeur auprès de Damavan Imaging SA et ingénieur à l'UTT (France)

OBJECTIF

Développer un système d'authentification biométrique, basé sur le réseau veineux, afin de pouvoir garantir le suivi de patients en contexte de crise humanitaire où les informations relatives à l'identité des victimes sont lacunaires ou même inexistantes.

3A. MEASUREME

ÉQUIPE

- Magali Leyvraz : doctorante à l'Institut de médecine sociale et préventive du CHUV (Suisse)
- Gustavo Rios : BSc et MSc en génie mécanique à l'EPFL (Suisse)
- Eric Nguyen Van : BSc et MSc en génie mécanique à l'EPFL (Suisse)

OBJECTIF

Développer une application mobile permettant de déterminer le poids d'un enfant et son état de malnutrition au moyen d'images, dans le but de faciliter le travail des soignants in-situ, ainsi que de collecter des banques de données sur la malnutrition, permettant la détection précoce de crise alimentaire, sans devoir entrer les données manuellement.

3A. DENTAL SIMULATOR

ÉQUIPE

- Professeur Dr Leandro Pereira : chirurgien-dentiste au São Leopoldo Mandic Dental School (Brésil)
- Rodrigo Dias Takase : software coordinator à Educational Technology Laboratory at University of Campinas (Brésil)
- Matheus Rocco : étudiant en informatique à University of Campinas (Brésil)

OBJECTIF

Développer une application mobile permettant aux étudiants en médecine dentaire de simuler dans un environnement virtuel des interventions telles que anesthésies, traitements de racine, poses d'implants, etc.

L'Idiap rencontre les ingénieur-e-s de demain !

Parce que les enfants d'aujourd'hui seront les ingénieurs – et les ingénieures ! – de demain, l'Idiap a à cœur de présenter son domaine d'activité, ainsi que les métiers qui y sont rattachés, au jeune public.

Hérisson sous gazon

A l'occasion du festival Hérisson sous gazon, qui a eu lieu à Charrat les 18 et 19 juin 2016, Hugues Salamin et Florent Monay ont mis sur pied un atelier intitulé « L'école des robots ». Le but ? Montrer à des enfants âgés de 6 à 14 ans comment apprennent les robots et leur présenter les concepts de programmation, d'apprentissage automatique et d'intelligence artificielle. Et ainsi présenter les activités et les compétences de l'Idiap à ce public particulier. Les robots devaient ici apprendre à reconnaître des bruits : les groupes de visiteurs se sont succédé tout au long du samedi, au son de leurs imitations des cris du phoque et autres.

« Oser tous les métiers »/« Futur en tous genres »
La journée « Oser tous les métiers »/« Futur en tous genres » est une journée croisée père-fille/mère-fils visant le décloisonnement des rôles et des métiers entre les sexes. Elle vise principalement à permettre aux filles et aux garçons d'explorer des univers professionnels nouveaux et d'élargir leurs horizons dans la perspective des choix à venir, sans se limiter aux secteurs traditionnellement féminins pour les filles ou masculins pour les garçons. Une thématique fort à propos dans le domaine d'activité, très masculin, de l'Idiap.

Lors de cette journée, organisée par le Secrétariat valaisan à l'égalité et à la famille, l'institut a accueilli, le 10 novembre 2016, 22 enfants, dont 3 garçons, fréquentant la 8e année Harmos à découvrir la palette de métiers qu'il abrite : du chercheur au gestionnaire de systèmes informatiques, du spécialiste de développement en programmation au responsable des projets européens. Après une présentation de l'Idiap et de ses métiers, les enfants ont pu retrousser leurs manches en participant à quatre ateliers d'une heure et demie chacun. Au programme : la salle des serveurs et la showroom de l'Idiap, le robot Nao, les Lego WeDo et enfin un atelier de traitement d'image en temps réel. Une expérience très positive que l'Idiap renouvelle depuis 2012 et avec laquelle il espère susciter de nouvelles vocations d'ingénieurs.

visages

38 — 43

Dribbler avec les chiffres

Arrivé à l'Idiap en 2004, Ed Gregg est aujourd'hui responsable des finances de l'institut. Mais c'est en tant que sportif professionnel que cet Américain a découvert la Suisse.

Son arrivée en Europe, Ed Gregg la doit à la fois un peu à sa grande taille – 2,13 mètres - et surtout à ses qualités de basketteur. C'est en effet en tant que joueur professionnel qu'il débarque sur le Vieux Continent en 1976. Il sillonne les parquets au sein de nombreuses formations, avec lesquelles il participe par deux fois à la Coupe d'Europe, avant de venir jouer en Suisse puis d'y entraîner plusieurs équipes.

A l'Idiap, Ed Gregg est le responsable des finances. Un domaine que cet Américain d'origine, âgé aujourd'hui de 63 ans, a étudié à la Utah State University et qu'il a décidé de retrouver à l'heure de sa reconversion professionnelle. Alors entraîneur au Tessin, il y intègre une fiduciaire avant de rejoindre son fils à Martigny, douze ans plus tard. A la recherche d'un emploi, il découvre l'Idiap. Au moment où l'Institut se lance dans les projets européens, sa langue maternelle convainc Pierre Dal Pont, ancien responsable des finances, de l'engager. D'abord à 20%, puis 30, 40 et ainsi de suite, jusqu'à reprendre le flambeau en 2010. Il y découvre une réalité financière très différente du secteur privé, mais se plaît dans le domaine de la comptabilité académique en raison de la très grande diversité des cas à traiter. Petit à petit il trouve sa place dans un monde de la recherche aussi étrange et nouveau que passionnant.

Ses nouvelles tâches ne l'ont pas pour autant détourné de sa passion pour le basket. Joueur, entraîneur, président, Ed Gregg a tenu tous les rôles du Martigny Basket, l'un des plus grands mouvements de jeunesse du canton. Aujourd'hui, il se consacre à la formation des plus petits pour leur plus grand plaisir.

3 étudiants du programme

Comprendre ce et ceux qui nous entourent

Mathématiques et psychologie : deux domaines scientifiques qu'il n'est guère courant de voir se côtoyer sur un curriculum vitae. Pourtant, c'est bien ainsi qu'ils figurent sur celui de Sandrine Tornay, doctorante à l'Idiap depuis août 2016.

« Les mathématiques ont toujours été pour moi une évidence, explique Sandrine Tornay. S'ils sont un langage pour comprendre le monde, la psychologie est un outil pour comprendre les autres. »

Après une maturité au collège de l'Abbaye de St-Maurice, la jeune chercheuse se décide pour les mathématiques, plus difficiles à appréhender par soi-même, et s'inscrit à l'EPFL. Son cursus lui impose la physique, pour laquelle elle n'a guère d'intérêt. Elle rejoint donc l'Université de Fribourg après la première année et ajoute les neurosciences et la psychologie à sa formation.

Sandrine Tornay est également active en politique. En 2011, elle est candidate aux élections fédérales des jeunes PDC. « Mes chances étaient nulles, mais j'avais envie de vivre ce processus de l'intérieur. » C'est au hasard d'une discussion politique avec Jean-Albert Ferrez qu'elle croise l'Idiap. « Je lui ai demandé si mon cursus de mathématicienne pouvait intéresser l'institut. Il m'a répondu que oui, et j'y ai effectué un stage durant l'été 2012, puis mon travail de master, tous deux sur le thème de la reconnaissance de la parole. »

Le sésame universitaire en poche, elle décide d'enseigner les mathématiques, à l'Ecole professionnelle technique et des métiers de Sion. Mais, ayant encore envie d'apprendre, elle reprend le chemin de l'Idiap et y entame un doctorat. Toujours sur le langage, mais des signes cette fois.

« En collaboration avec l'Interkantonale Hochschule für Heilpädagogik de Zurich et l'Université de Surrey (UK), nous développons un système d'évaluation automatique du langage des signes suisse allemand (DSGS) : lorsqu'un apprenant effectue des signes devant la caméra, la machine l'évalue et l'aide en les commentant. »

Des robots et des montagnes

Lorsque l'on a grandi entouré de montagnes et qu'on adore les escalader, les parcourir en ski de randonnée ou les dévaler en VTT, la localisation de l'Idiap est idéale. Et même si ce n'est pas la raison principale qui a poussé Emmanuel Pignat à ouvrir les portes de l'institut de Martigny, la perspective de pouvoir travailler à proximité des cimes a quelque chose de très motivant.

Depuis mars 2016, ce natif de Collombey-Muraz travaille comme doctorant au sein du groupe « Robot Learning & Interaction » de Sylvain Calinon, dans le cadre du projet de recherche européen i-Dress (voir article p.26). Un travail qui est en fait la continuation du travail de master effectué à l'Idiap à la fin de ses études en microtechnique à l'EPFL.

« La robotique est un domaine qui me passionne. Avoir un institut de recherche tel que l'Idiap dans une région qui réunit à la fois mes intérêts professionnels, mes hobbies et ma famille, c'est extraordinaire. C'est donc dans l'idée d'y faire un doctorat que j'ai décidé de l'intégrer dans le cadre de mon master », explique Emmanuel Pignat.

Une année après le début de sa thèse, les travaux avancent bien : « Aujourd'hui, Baxter arrive à enfiler la veste d'un patient », poursuit-il. Et de s'attaquer maintenant à une tâche plus complexe encore : lui faire enfiler une paire de chaussures.

Entre biologie et ingénierie

Il arrive parfois que l'agenda d'un candidat intéressant ne corresponde pas exactement à celui d'un projet de recherche. C'était le cas pour Adrian Shajkofci à la fin de son master. Le programme Valais-Wallis Ambition a permis de résoudre le problème à la satisfaction de tous.

Pianiste à ses heures, Adrian Shajkofci aime bidouiller les sons. Une activité artistique pas si éloignée des recherches qui occupent ce doctorant de l'Idiap depuis septembre 2016. En effet, dans le groupe « Computational Bioimaging » de Michael Liebling, il traite des images, plus précisément des images microscopiques d'embryons de poisson. « Ces cœurs battent trois à cinq fois par seconde, explique le chercheur. Visualiser ce phénomène à des fins de recherche, par exemple sur les malformations cardiaques, impose une excellente résolution spatiale et temporelle. » Doper ces images nécessite de travailler à la fois sur l'acquisition et le traitement du signal. A cette fin, Adrian Shajkofci travaille sur le développement d'une plateforme de microscopie automatisée recourant à des techniques d'intelligence artificielle et d'apprentissage automatique. Titulaire d'un bachelor en sciences de la vie et d'un master en bio-ingénierie, tous deux réalisés à l'EPFL, Adrian Shajkofci est à l'aise dans son rôle de pivot au sein d'une équipe interdisciplinaire réunissant des biologistes et des ingénieurs. « Ce projet est un mélange idéal de ce que je sais faire et de ce que j'aime faire, dans le cadre agréable d'un institut de taille humaine », conclut-il.

Prix Idiap

Cijo Jose

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 2016, le prix de la recherche a été attribué à Darshan Santani pour son autonomie et sa motivation exceptionnelle dans ses recherches portant sur la compréhension de l'utilisation des espaces urbains à partir de données géolocalisées issues des réseaux sociaux tels que Twitter et Foursquare. Celui de la publication a été remis à Cijo Jose pour son excellent article scientifique intitulé « Scalable Metric Learning via Weighted Approximate Rank Component Analysis ».

Darshan Santani

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.

N. LE, D. WU, S. MEIGNIER ET J-M.ODOBEZ.
MEDIAEVAL 2016,HILVERSUM, SEPTEMBER 2016,
WINNER OF THE PERSON DISCOVERY CHALLENGE
EUMSSI team at the MediaEval Person Discovery Challenge 2016

NIKOLAOS PAPPAS, M. TOPKARA, M. REDI, B. JOU,
T. CHEN, H. LIU, & S.-F. CHANG
ACM INTERNATIONAL CONFERENCE ON MULTIMEDIA RETRIEVAL (ICMR), NEW YORK, JUNE 2016, BEST MULTIMODAL PAPER AWARD
Multilingual Visual Sentiment Concept Matching

Thèses achevées

Six étudiants ont achevé leur thèse de doctorat en 2016 : Rémi Lebret, Joël Legrand, Dimitri Palaz, Nikolaos Pappas, Darshan Santani, Raphaël Ullmann.

BUILDING WORD EMBEDDINGS FOR SOLVING NATURAL LANGUAGE PROCESSING
Rémi Philippe LEBRET
26 septembre 2016
Directeurs de thèse : Prof. Hervé Bourlard et Dr Ronan Collobert
Membres du jury : Prof. Jean-Philippe Thiran, Dr Jean-Cédric Chappelier, Dr Yves Grandvalet et Prof. Giuseppe Attardi

WORD SEQUENCE MODELING USING DEEP LEARNING : AN END-TO-END APPROACH AND ITS APPLICATIONS
Joël Yvon Roland LEGRAND
28 octobre 2016
Directeurs de thèse : Prof. Hervé Bourlard et Dr Ronan Collobert
Membres du jury : Dr Denis Gillet, Dr Jean-Marc Vesin, Dr Claire Gardent et Dr Cho Kyunghyun

TOWARDS END-TO-END SPEECH RECOGNITION
Dimitri Palaz
13 juillet 2016
Directeurs de thèse : Prof. Hervé Bourlard et Dr Ronan Collobert
Membres du jury : Dr Jean-Marc Vesin, Prof. Jean-Philippe Thiran, Prof. Steve Renals et Dr Samy Bengio

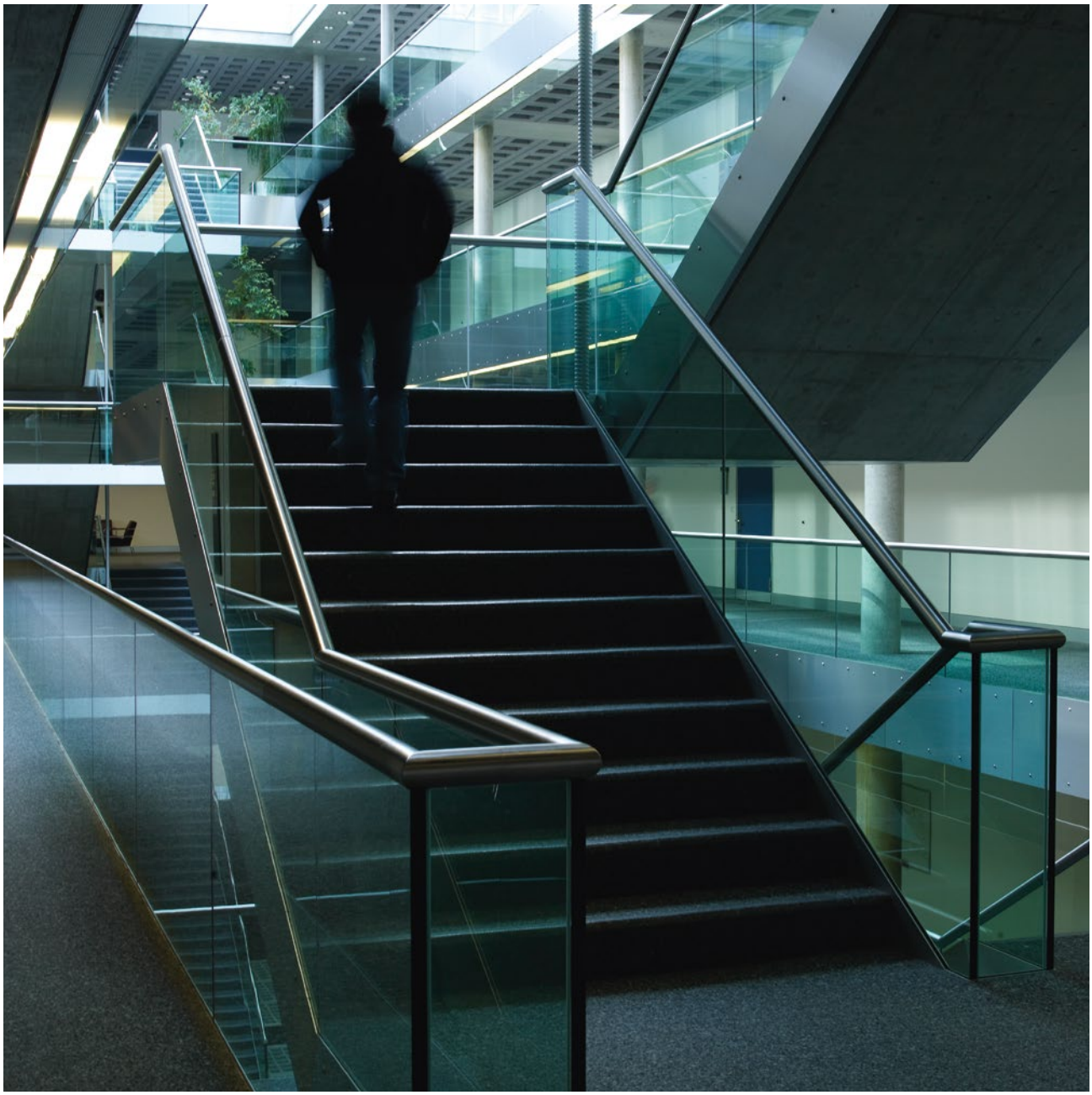
LEARNING EXPLAINABLE USER SENTIMENT AND PREFERENCES FOR INFORMATION FILTERING
Nikolaos Pappas
4 mars 2016
Directeurs de thèse : Prof. Hervé Bourlard et Dr Andrei Popescu-Belis
Membres du jury : Prof. Sabine Süsstrunk, Dr Jean-Cédric Chappelier, Prof. Thomas Hofmann et Prof. Shih-Fu Chang

COMPUTATIONAL ANALYSIS OF URBAN PLACES USING MOBILE CROWDSENSING
Darshan Santani
17 novembre 2016
Directeur de thèse : Prof. Daniel Gatica-Perez
Membres du jury : Prof. Pascal Frossard, Prof. Kristof Van Laerhoven et Dr Andres Monroy-Hernandez

CAN YOU HEAR ME NOW ? AUTOMATIC ASSESSMENT OF BACKGROUND NOISE INTRUSIVE-NESS AND SPEECH INTELLIGIBILITY IN TELECOMMUNICATIONS
Raphaël Marc ULLMANN
14 juin 2016
Directeur de thèse : Prof. Hervé Bourlard
Membres du jury : Prof. Jean-Philippe Thiran, Dr Hervé Lissek, Prof. Martin Cooke et Dr John Beerends

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

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



scientific inserts

I — XLIV

1. Speech and Audio Processing

OVERVIEW

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

GROUP OVERVIEW

Speech processing has always 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, and automatic speech intelligibility evaluation. The expertise and activities of the group encompass statistical automatic speech recognition (based on Hidden Markov Models – HMM, or hybrid systems exploiting connectionist approaches and Deep Neural Networks – DNN), text-to-speech (TTS), speaker recognition (with extensions towards forensics) 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 2016 was composed of 1 head of group, 3 principal investigators, 3 research associates, 6 postdocs, 9 PhD students and 4 interns. In addition to that, the group also worked closely with two Senior Development Engineers from the development team at Idiap on industrial research projects.

KEY SCIENTIFIC OUTPUTS

Idiap is always significantly contributing 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 " deep " neural networks (DNNs) technology, and the associated Kaldi Automatic Speech Recognition (ASR) toolkit, now used by most of the international speech community. Several key contributions were achieved in speaker recognition tasks, particularly focusing on textdependent speaker verification scenarios and information fusion for large-scale speaker identification. More recently, the group 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. Much of the group's recognition has come from contributions to the speech and audio community. These include the AMI corpus[1] used for multiparty conversational speech processing, the Juicer ASR decoder[2], and major contributions to the Kaldi[3] open-source ASR toolkit. More recently, the MediaParl[4] and Walliserdeutsch[5] databases have been released ; many other contributions are on the GitHub site[6]. This highlights that, although the group has traditionally worked with English speech, many recent projects are multi-lingual in nature. This has in turn become a focus of the group ; one on which we are well placed to capitalise given our geographical location.

1 www.amiproject.org
2 juicer.amiproject.org/juicer
3 kaldi-asr.org
4 www.idiap.ch/dataset/mediaparl
5 www.idiap.ch/dataset/walliserdeutsch
6 github.com/idiap

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 and German, Idiap is actively carrying out research in several ASR directions, including:

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

- **Cross-lingual and multi-lingual speech recognition**
  Within the scope of the SUMMA project (Figure 4), 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.

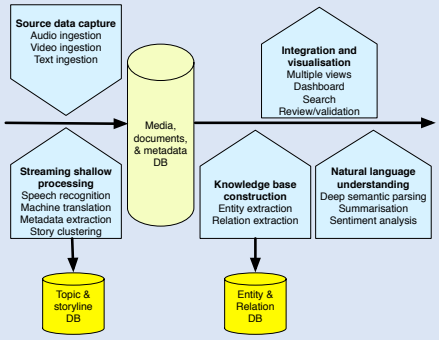


Figure 4

An overview of the SUMMA project, a current focus for multilingual ASR 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 new collaboration in 2016 with Swisscom has enabled us to extend our Swiss German work.

- **Template-based ASR**
  Idiap continues to work on template-based or exemplar-based ASR using posterior-based features in the context of projects like SNSF projects A-MUSE, PHASER, and PHASER-QUAD, and CTI project MultiVEO[7]. 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.

[7] www.idiap.ch/en/scientific-research/projects/MULTIVEO

- **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).

- **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.

- **Failure analysis of ASR and HSR decoding channels**
  This is a novel strategy developed in 2016 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 sub-word components from their probabilistic representation. Information theoretic measures are developed to assess and quantify the information loss in acoustic modeling for ASR decoding using hidden Markov models.

SPEECH SYNTHESIS AND CODING

- **Text-to-speech synthesis (TTS)**
  Although newer than ASR, TTS is now an established venture for the speech group at Idiap. TTS has been central to several projects, including: SIWIS[8] (Spoken Interaction With Interpretation in Switzerland), D-BOX[9] and SP2[10] (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. A key theme of SIWIS is adaptation of speech prosody, i.e., spoken emotion. This will ensure that, e.g., questions and stressed phrases remain as such after translation. To this end, we have developed novel models of speech prosody. The goal of SIWIS was mirrored in the SP2 project with a focus on eastern-European languages.

- **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.

---

[8] www.idiap.ch/project/siwis

[9] www.idiap.ch/en/scientific-research/projects/DBOX

[10] www.idiap.ch/en/scientific-research/projects/SP2

SPEAKER RECOGNITION AND SPEECH ANALYTICS

In the context of the European SIIP project[11], illustrated by Figure 5, 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 a proof-of-concept event, 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.

As discussed in Section 2.8, and as part of the SNSF Project UniTS[12], 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.

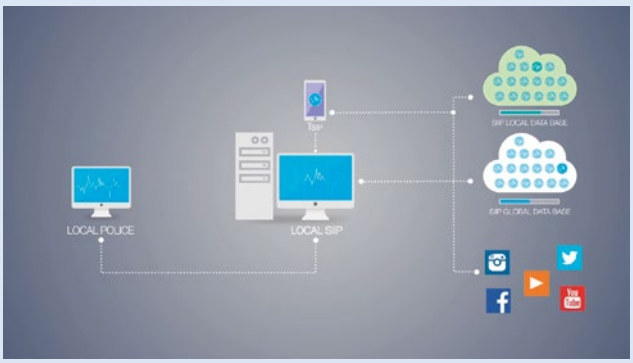


Figure 5

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 2016, Idiap became active in audio forensics which is a field of forensic science for extracting 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

In 2016, Idiap became active in audio forensics which is a field of forensic science for extracting 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.

- **Query-by-example spoken term detection (QbE-STD) based on subspace modelling and detection**
  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 characterization and detection of the low-dimensional subspace of the DNN phonetic posteriors.

- **Large scale spoken query indexing**
  In 2016, Idiap also developed a powerful methodology for largescale spoken query retrieval relying on hashing.

[11] www.siip.eu

[12] www.idiap.ch/en/scientific-research/projects/UNITS

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. Specifically, leveraging from our posterior feature based ASR research, we have developed a novel approach where the intelligibility is objectively assessed by estimating a phonetic difference score between reference speech signal and test speech signal. A distinctive advantage of the proposed approach is that the reference speech signal can be replaced by a text transcription of the test speech signal and word level intelligibility assessment can be effectively carried out. In 2016, as part of continuation of armasuisse project OMSI-2015_ARMASUISSE[13], our R&D mainly focused on intelligibility assessment in realistic conditions, such as field operations, narrow band radio transmissions, analog versus digital radio transmissions.

SIGN LANGUAGE RECOGNITION AND ASSESSMENT

In the context of SNSF Sinergia project SMILE[14], Idiap has initiated 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 2016, our research activities in the area of microphone array based speech processing and speaker diarization were mainly addressing the problem source localization and reconstruction through binary sparse coding framework.

KEY PUBLICATIONS

[1] A. Asaei, M. Taghizadeh, S. Haghighatshoar, B. Raj, H. Bourlard, V. Cevher, " Binary Sparse Coding of Convolutive Mixtures for Sound Localization and Separation via Spatialization ", *IEEE Transactions on Signal Processing,* Vol. 64, N. 3, pages 567–579, 2016.

[2] M. Cernak, A. Lazaridis, A. Asaei, P. Garner, " Composition of Deep and Spiking Neural Networks for Very Low Bit Rate Speech Coding ", *IEEE/ACM Trans. on Audio, Speech and Language Processing,* pages 2301–2312, 2016.

[3] M. Razavi, R. Rasipuram, M. Magimai.-Doss, " Acoustic data-driven grapheme-to-phoneme conversion in the probabilistic lexical modeling framework ", *Speech Communication,* Vol. 80, pages 1–21, 2016.

[4] I. Himawan, P. Motlicek, D. Imseng, S. Sridharan, " Feature mapping using far-field microphones for distant speech recognition ", *Speech Communication,* Vol. 83, pages 1–9, 2016.

[5] M. Ferràs, S. Madikeri, and H. Bourlard, " Speaker Diarization and Linking of Meeting data ", *IEEE Trans. on Audio, Speech and Language Processing,* 24(11), pages 1935–1945, 2016.

[13] www.idiap.ch/en/scientific-research/projects/OMSI_EXT

[14] www.idiap.ch/en/scientific-research/projects/SMILE

2. Computer Vision and 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

The main objective of the Computer Vision and Learning group is the development of novel machine learning and statistical methods, with an emphasis on their algorithmic efficiency. The application domain we focus on encompasses any processing of images and video streams, such as object detection and scene analysis, tracking of objects and biological structures, and image recognition in general.

The group was composed in 2016 of one head of group, two post-doctoral researchers, four PhD students, two visiting PhD students, one intern, and it had strong interactions with the software development team. Francois 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

Our work has resulted in contributions in four different domains: adaptive sampling to deal with very large training sets, large-scale state-of-the-art clustering using geometrical bounds, large-scale metric learning with an orthogonal-inducing regularizer that has a great potential for other learning problems, and mean-field inference combined with deep learning feature extraction for multi-camera object detection.

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

MACHINE LEARNING

- **Efficient 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 insures that it is not necessary.

  In the context of the MASH2 project, 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 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 batch 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**

  In the SNSF project DASH, we were interested in speedingup " hard sample " extraction. Most of the state-of-the-art detection algorithms are trained in a sequential manner, improving iteratively the detector by training it with samples miss-classified by its previous version. Since detectors are very accurate, finding a sufficient number of hard samples is computationally intensive, and has traditionally a cost proportional to the number of images to visit.

  Our approach consists in using the prior structure provided for any large collection of images 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 practice, our approach alternates between looking into images it has poor information about, and images from which it has observed a high success rate in finding hard samples.

- **High-dimension similarity measures**

  The WILDTRACK project is 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 fields 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 person.

The technique we have developed casts this problem as a rank optimization 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.

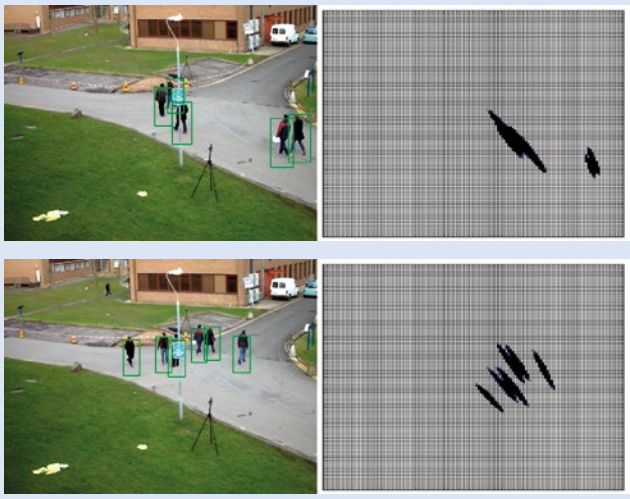


Figure 6

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

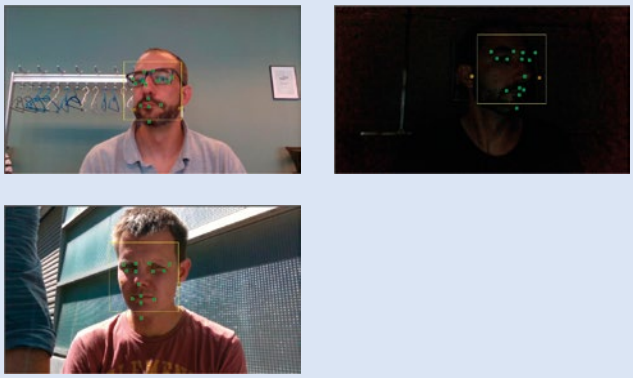


Figure 7

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

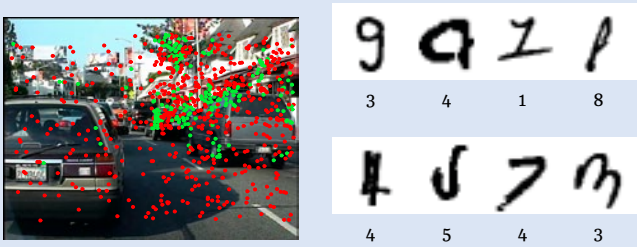


Figure 8

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).

IMAGE ANALYSIS

- **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 scene which 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.

- **Manufacturing quality control with deep neural networks**

  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.

  To address the problem, we are investigating generative deep network able to learn from the proper examples alone a sound criterion to detect failures. Our current models are able to decompose an image into two components, one standing for the background image of the materials, and the other for the shape of the manufactured component, which is the only one used to assess the quality of the process.

- **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 drastically the performance, in particular when dealing with poor lighting conditions.

- **Deep learning for multi-camera detection**
  We have a long-lasting collaboration with the CVLab at EPFL around the design of a robust multi-camera tracking system, now in the context of the SNSF WILDTRACK project, in collaboration with ETHZ. We have in particular improved our stateof- the-art mean-field inference scheme Probabilistic Occupancy Map, to leverage deep learning methods for the image feature extraction.

  In parallel, we have developed at Idiap a new approach to adapt a monocular deep learning detector to a multi-camera context. We 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.

- **Advertisement substitution in sport events**
  In a collaboration with ES Concept SA, we have developed a pipeline of algorithms to perform advertisement substitution on-the fly in broadcast video streams. The objective is to deliver a solution able to automatically and accurately localize perimeter advertisement during a sport event, and replace it with another one.

  We have implemented a pipeline which demonstrates such a system, in particular the stability of the estimated board positions in the image plan.

KEY PUBLICATIONS

[1] J. Newling and F. Fleuret. Fast mini-batch k-means by nesting. In Proceedings of the international conference on Neural Information Processing Systems (NIPS), pages 1352–1360, 2016.

[2] C. Jose and F. Fleuret. Scalable Metric Learning via Weighted Approximate Rank Component Analysis. In Proceedings of the European Conference on Computer Vision (ECCV), pages 875–890, 2016.

[3] L. Lefakis and F. Fleuret. Jointly Informative Feature Selection Made Tractable by Gaussian Modeling. Journal of Machine Learning Research (JMLR), 17(182):1–39, 2016.

[4] O. Canévet and F. Fleuret. Large Scale Hard Sample Mining with Monte Carlo Tree Search. In Proceedings of the IEEE international conference on Computer Vision and Pattern Recognition (CVPR), pages 5128–5137, 2016.

[5] X. Wang, E. Turetken, F. Fleuret, and P. Fua. Tracking Interacting Objects Using Intertwined Flows. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 38(11): 2312–2326, 2016.

3. Social Computing

OVERVIEW

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

GROUP OVERVIEW

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

The Social Computing group in 2016 was composed of one group head, one research associate, four postdoctoral researchers, four PhD students, and one intern. Research lines investigated in 2016 included: social media analytics and mobile crowdsourcing for cities, 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 document youth nightlife patterns and to understand urban perception in cities; (2) multimodal analysis and inference of social constructs in face-to-face employment interviews; and (3) Maya hieroglyphic categorization and retrieval integrating crowdsourcing for data labeling and automatic shape analysis. One EPFL PhD student graduated in 2016 (12 EPFL PhD students graduated since 2002.)

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

SOCIAL MEDIA ANALYTICS AND MOBILE CROWDSOURCING IN CITIES

Our work in this domain spans several research lines. First, as part of the SNSF Youth@Night project[15] (A 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 and online crowdsourcing and social media analytics (Twitter, Instagram, Foursquare) to characterize urban phenomena at scale in cities, both in Switzerland and other world regions[16]. This included the automatic characterization of urban areas according to patterns of youth night activity in a study done for the City of Lausanne[17]; the automatic characterization of physical ambiance from night videos crowdsourced by youth in Zurich and Lausanne (see [1] at the end of this section); and the recognition of social ambiance from social media images of indoor places using deep learning techniques [2] (Figure 9).

In the context of the Civique social innovation platform enabled by the Valais+ project (supported by Loterie Romande)[18], we collaborated with the City of Lausanne to design and implement a mobile survey to collect citizen experiences about street harassment. The results of the study showed the prevalence of this phenomenon, received wide media coverage, and has resulted in concrete city action[19]. This highlights the importance of working jointly with cities to generate insights from data with the participation of city government and citizens. Several other collaborations using the Civique platform are under development.

[15] www.youth-night.ch

[16] www.youtube.com/watch?v=71ht15VAoLw

[17] www.youth-night.ch/static/publications/yatn_rr83.pdf

[18] www.civique.org

[19] www.idiap.ch/en/allnews/results-of-street-harassment-survey-presented-in-lausanne

[20] www.idiap.ch/project/ubimpressed

UBIQUITOUS INTERACTION ANALYTICS

In the context of the SNSF UBImpressed interdisciplinary project[20] (Ubiquitous First Impressions and Ubiquitous Awareness), we have developed computational methods to analyze dyadic interactions in the workplace using sensors (cameras, Kinect devices, microphone arrays, smartphones, and wearables), and automatically infer key social variables (Figure 10). This is joint work with the University of Lausanne and Cornell University. Specifically, we have investigated connections between nonverbal behavior and hirability and other related variables. In all cases, nonverbal cues are automatically measured from audio and video streams, including speaking activity, prosody, and body motion [3].

Figure 10

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

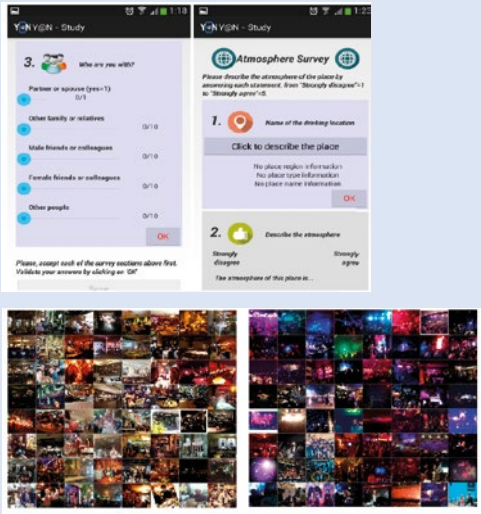


Figure 9

Left: Youth@Night survey logger app. Right: Images classified as Restaurant and Stageby deep network.

This line of work is complemented by the analysis of online video introductions (supported by a project funded by The Ark), where we investigated the issue of interpersonal judgment of dimensions related to impressions of hirability, as recently shown in our work [4].

VISUAL ANALYSIS OF MAYA HIEROGLYPHS

In the context of the SNSF MAAYA project[21] (Multimedia Analysis and Access for Documentation and Decipherment of Maya Epigraphy), we have developed methods for automatic visual analysis of hieroglyphs for the three remaining ancient Maya codices located in museums in Europe (Dresden, Paris, Madrid). Our most recent work has investigated methods for Maya hieroglyphic categorization and retrieval using sparse autoencoders [5] and for glyph stroke segmentation (Figure 11); the use of online crowdsourcing for non-expert glyph labeling, namely localization and segmentation of single signs in glyph-blocks; and graph-based interactive visualization of glyphs according to visual similarity[22]. The visualization work was evaluated with the participation of the general public at two events in 2016: the Idiap Innovation Day (September), and the EPFL Open Days (November).

[21] www.idiap.ch/project/maaya

4. Perception and Activity Understanding

Figure 11

Glyph stroke segmentation, using 500 or 3000 superpixels as spatial support, and two segmentation models (appearance-based method, and when a Conditional Random Field is added).

KEY PUBLICATIONS

[1] J. D. Santani, J.-I. Biel, F. Labhart, J. Truong, S. Landolt, E. Kunstche, and D. Gatica-Perez, The Night is Young: Urban Crowdsourcing of Nightlife Patterns, in Proc. ACM Int. Joint Conf. on Ubiquitous and Pervasive Computing (Ubicomp), Heidelberg, Sep. 2016.

[2] D. Santani, R. Hu, and D. Gatica-Perez, InnerView: Learning Place Ambiance from Social Media Images, in Proc. ACM Int. Conf. on Multimedia (MM), Amsterdam, Oct. 2016.

[3] S. Muralidhar, L. Nguyen, D. Frauendorfer, J.-M. Odobez, M. Schmid Mast, and D. Gatica-Perez, "Training on the Job: Behavioral Analysis of Job Interviews in Hospitality," in Proc. ACM Int. Conf. on Multimodal Interaction (ICMI), Tokyo, Nov. 2016.

[4] L. Nguyen and D. Gatica-Perez, "Hirability in the Wild: Analysis of Online Conversational Video Resumes," IEEE Trans. on Multimedia, Vol. 18, No. 7, pp. 1422–1437, Jul. 2016.

[5] G. Can, J.-M. Odobez, and D. Gatica-Perez, "Evaluating Shape Representations for Maya Glyph Classification," ACM Journal on Computing and Cultural Heritage, Vol. 9, Issue 3, Nov. 2016.

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 2016, the group was composed of one group head, two post-doctoral researchers, three PhD students, one intern, and focused collaborations with the software team.

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, the PAU team ranked first at the 2016 MediaEval Person discovery challenge. In 2016, its recent 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

[22] lab.idiap.ch/maaya/

work, the group has investigated deep learning methods for several tasks like gesture recognition, audio-visual speaking activity modeling, gaze, or shape recognition. It has also worked on the integration of its sensing technology into realtime perceptual systems for human-robot interaction (Pepper platform, EU MuMMER project). During the period 2012–2016, the group published 16 journal papers and around 45 conference papers.

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

DEEP LEARNING

In the past years, the team has devoted substantial efforts in investigating the deep learning paradigm to address PAU related tasks. In [2] we studied a multimodal Deep Dynamic Neural Networks (DDNN) framework for the segmentation and recognition of short spontaneous communicative gestures, as illustrated in Fig. 12a. We adopted a semi-supervised hierarchical approach based on Hidden Markov Models learning and fusing emission probabilities from high-level spatio-temporal representations trained using deep neural networks suited to the input modalities: a Gaussian-Bernoulli Deep Belief Network (DBN) to handle the dynamics of skeletal joint information (Fig. 12b), and a 3D Convolutional Neural Network (3DCNN) to manage and fuse batches of depth and RGB images (Fig. 12c). We investigated in [5] deep-temporal representations for the modeling of audio-visual speaking activities. More precisely, to differentiate dubbing from genuine talking situations, we proposed a method relying on canonical correlation analysis to learn a joint multimodal space, and on long short term memory (LSTM) networks to model cross-modality temporal dependencies linked to anticipation and retention phenomena between the visual and audio streams. The method achieved promising results on challenging data from the EU EUMSSI project.

Other currently addressed tasks include shape representations of Maya Glyph (with Social computing group, see Fig. 11), body-parts localization, gaze estimation, or sound source localization in HRI.

MULTIPLE OBJECT TRACKING

Our previous work resulted in an enhanced Conditional Random Field tracking-by-detection framework with important characteristics making it successful 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 avoiding tedious annotations. In 2016, we extended this framework for face track extraction in broadcast data. Thanks to the frame sub-sampling capacity and the learning of program dependent association costs, the model performed better than previous methods with a faster processing, which is useful when processing large video repositories.

GAZE, NON-NERBAL BEHAVIOR EXTRACTION, HUMAN-ROBOT INTERACTIONS

Our long-standing work on NVB perception has been extended in several ways. In the context of the Swiss UBImpressed project, we have worked on robust and accurate head pose tracking from RGB-D data. The approach combines the benefits of the online fitting of a 3D face morphable model with the online 3D reconstruction of the full head. This provides more support when handling extreme head poses, and permits to handle well difficult UBImpressed sequences of role play where a student from the Hospitality industry has to handle a difficult client at a registration desk, as illustrated in Fig. 14. Our long-term research on gaze estimation from cheap RGB-D (depth) cameras like Kinect [1] has been pursued by improving the robustness of our G3E model, which leverages the benefit of the estimation of a geometric model of the eye with the robustness of image driven appearance-based techniques. Figure 13 illustrates the result of applying our tools for gaze coding in multiparty situations.

A: B: C:

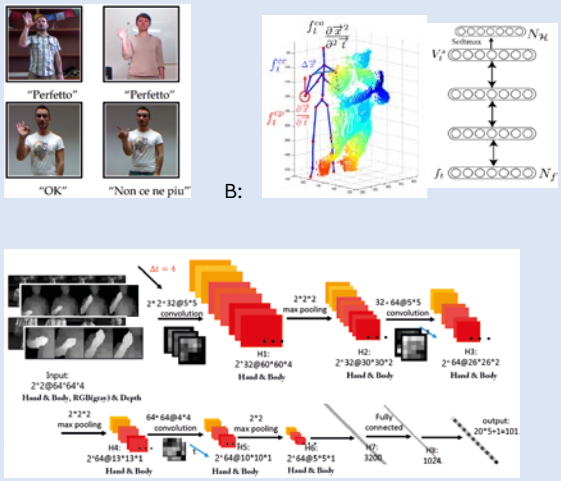


Figure 12

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

In the context of the EU MuMMER project, we have integrated our multi-person tracker on the Pepper robotics platform. This include the extraction of non-verbal cues (attention, head gestures) from both visual only and RGB-D data, and will allow to benchmark our methods in real interactions. In collaboration with KTH, we have worked on the perception of attentiveness from audio-visual (AV) feedbacks [3], investigating which features of head nods influences attentiveness perception, or whether visual or verbal backchannels or their combination are perceived to be more attentive. Such analysis is useful for endowing robots with the capacity to demonstrate variable level of attentiveness thanks to the synthesis of appropriate subtle AV feedback tokens.

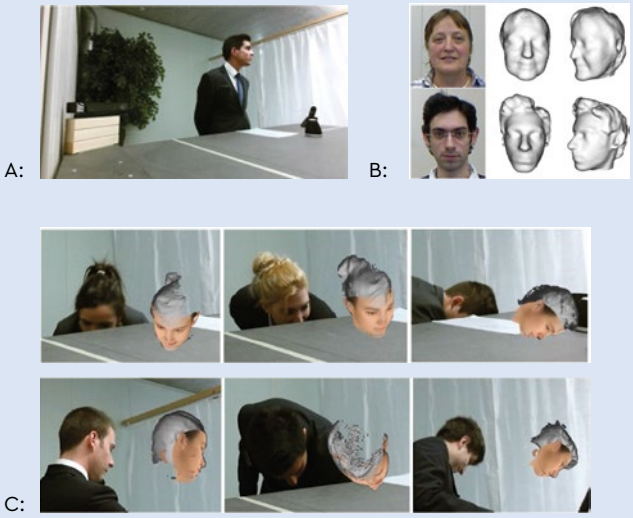


Figure 14

Robust head pose tracking. a) UBImpressed registration desk set-up. b) Example of automatically reconstructed heads. c) Example of tracking results in adverse conditions.

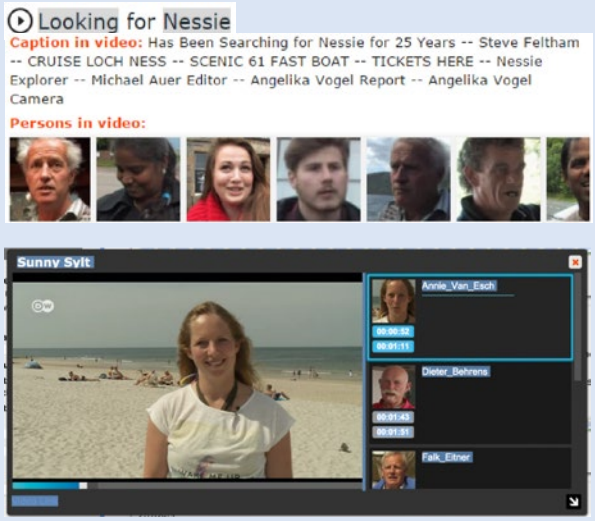


Figure 15

Person identification within the EUMSSI project. Left illustrating the found face clusters of a given Deutsche Welle video document (as well as automatically extracted OCR). Right: fast browsing of documents based on face occurrences.

MULTIMODAL FACE AND PERSON DIARIZATION AND NAMING

Identifying people appearing and speaking in multimedia data as we do 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 our broadcast programs, as illustrated in Fig. 15. However, person discovery in the absence of prior identity knowledge requires accurate association of visual and auditory cues which in broadcast data faces additional challenges due to narrated voices over muted scenes or dubbing in different languages. In 2016, we have thus investigated dubbing detection (see [5] and deep learning part above), and improved the robustness and efficiency of our video processing pipeline, with an emphasis on the precision of the information delivered to and visualized by the EUMSSI system users.

KEY PUBLICATIONS

[1] 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.

[2] 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.

[3] C. Oertel, J. Lopes, Y. Yu, K. Funes, J. Gustafson, A. Black, and J.-M. Odobez. Towards building an attentive artificial listener: On the perception of attentiveness in audio-visual feedback tokens. In Int. Conf. on Multimodal Interactions (ICMI), 2016.

[4] Paul Gay, Sylvain Meignier, Paul Deleglise, and Jean-Marc Odobez. CRF-based context modeling for person identification in broadcast videos. Frontiers in ICT: Computer Image Analysis, 3, 2016.

[5] N. Le and J.-M Odobez. Learning multimodal temporal representation for dubbing detection in broadcast media. In ACM Multimedia, Amsterdam, 2016.

5. Robot Learning and Interaction

OVERVIEW

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

GROUP OVERVIEW

The Robot Learning and Interaction group, created in 2014, focuses on humancentric robotic applications in which the robots can learn new skills by interacting with the end-users. From a machine learning perspective, the challenge is to acquire skills from only few demonstrations and interactions, with strong generalization demands. It requires: 1) the development of intuitive active learning interfaces to acquire meaningful demonstrations ; 2) the development of models that can exploit the structure and geometry of the acquired data in an efficient way ; 3) the development of adaptive control techniques that can exploit the learned task variations and coordination patterns.

The Robot Learning and Interaction group in 2016 was composed of one postdoctoral fellow, three PhD students, four MSc students and two BSc students.

KEY SCIENTIFIC OUTPUTS

Development of robot learning and control algorithms that can be applied to a wide range of applications, for robots that are either close to us (assistive robots in i-Dress), parts of us (prosthetic hands in TACT-HAND), or far away from us (manipulation skills in deep water with DexROV). Concerned about reproducible research, the group regularly releases open source codes accompanying 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

Figure 16

Human-centric robotic applications. Left: Teleoperation assistance for the control of a bimanual robot (DexROV project), where the robot progressively assists the teleoperator when repetitive or well structured tasks are achieved, as a form of human-robot collaboration. As an example, if the robot has observed 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 by delegating the orientation tracking aspect to the robot. Center: Personalized assistance in dressing (i-Dress project), where the robot learns by demonstration the preferences and types 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. Right: 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.

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. The collected datapoints are organized in a tensor form (multidimensional arrays). A statistical learning approach is then used to determine the variability and coordination patterns in the movement by considering the different frames of reference simultaneously. This approach allows the orchestration of the different frames to reproduce the movement in a new situation (typically, for adaptation of the 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.

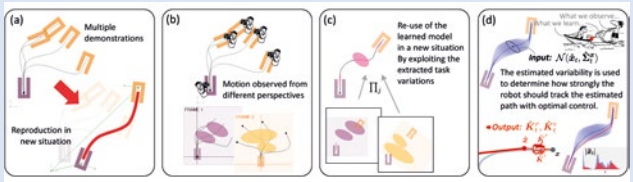


Figure 17

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

A DISTINCTIVE VIEW OF OPTIMAL CONTROL

Model predictive control (MPC) and linear quadratic tracking (LQT) are widespread techniques in robot control, but the core formulation of these problems and their 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 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 in a fast and simple manner, by considering variation and coordination constraints. Instead of learning the trajectory directly, the approach allows the learning of the underlying controller 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.

GEOMETRY-AWARE STATISTICAL LEARNING AND REGRESSION

Figure 18

Illustration of iterative gradient descent (left) and Gaussian mixture model (right) on a Riemannian manifold, where the centers of the Gaussians are defined on the manifold, and the covariances are defined in the tangent spaces of the corresponding centers.

Data in robot learning and control problems have structures and symmetries that are often underexploited. We explore the use of Riemannian manifolds to provide geometry-aware algorithms that can be exploited in several sensing and actuation problems. Examples of Riemannian manifolds in robotics with well known geometries include stiffness and damping gains, inertia, manipulability ellipsoids (symmetric positive definite matrices), orientations (unit quaternions), period movements (phase variable on circle manifold), or rotary joints (e.g., a 2-link planar robot forms a torus manifold).

We explore the use of differential geometry techniques to treat data of various forms in a uniform probabilistic manner (including data in standard Euclidean spaces). In particular, we exploit the potential that several Riemannian manifolds can be combined together to treat problems such as learning/controlling the pose of an end-effector (position and orientation), or learning/controlling both periodic and discrete movements.

It can typically be used to revisit common optimization problems in robotics formulated in standard Euclidean spaces, by removing the constraint of staying on a given manifold, and instead considering an unconstrained problems in spaces that inherently take into account this constraint.

KEY PUBLICATIONS

[1] Tanwani, A.K. and Calinon, S. (2016). Learning Robot Manipulation Tasks with Task-Parameterized Semi-Tied Hidden Semi-Markov Model. IEEE Robotics and Automation Letters (RA-L), 1:1, 235–242.

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

[3] Rozo, L., Calinon, S., Caldwell, D.G., Jimenez, P. and Torras, C. (2016). Learning Physical Collaborative Robot Behaviors from Human Demonstrations. IEEE Trans. on Robotics, 32:3, 513–527.

[4] Rozo, L., Silvério, J., Calinon, S. and Caldwell, D.G. (2016). Learning Controllers for Reactive and Proactive Behaviors in Human-Robot Collaboration. Frontiers in Robotics and AI, 3:30, 1–11.

[5] Billard, A., Calinon, S. and Dillmann, R. (2016). Learning From Humans. Siciliano, B. and Khatib, O. (eds.). Handbook of Robotics, 2nd Edition, Chapter 74, pp. 1995–2014. Springer.

6. Uncertainty Quantification and Optimal Design

OVERVIEW

Head: PD. Dr. David Ginsbourger (Ph.D. Mines Saint-Etienne 2009, Habilitation Univ. Bern 2014)

GROUP OVERVIEW

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

UQOD started at Idiap in September 2015, with members coming from and keeping strong academic ties to the Institute of Mathematical Statistics and Actuarial Science (IMSV) of the University of Bern (UniBE). During the year 2016, the UQOD group has been composed of a permanent senior researcher, a visiting and then Idiap postdoctoral researcher (who now took over a permanent position abroad), a long term visiting IMSV PhD student (now Idiap postdoc), an intern, and three UniBE students occasionally visiting the group (two master students and a PhD student in co-tutelle with Marseille).

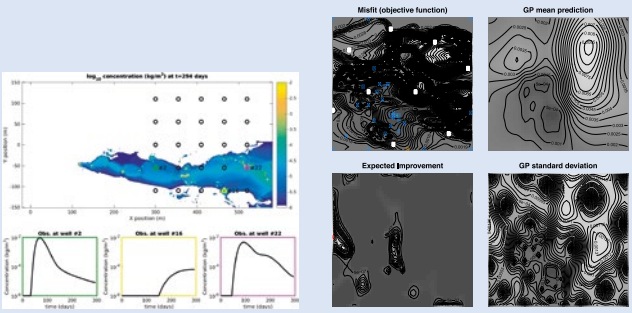


Figure 19

Bayesian optimization on a contamination source identification problem.

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 models. Other recent results deal with the interplay between the choice of covariance kernels and properties of Gaussian Process paths, with implications in high-dimensional Gaussian Process modelling and in function prediction under structural constraints. Ongoing work also encompasses novel algorithms for non-stationary modelling of extremes with application in climate sciences, as well as a collaboration with hydrogeologists towards flow simulation-based contaminant source localization.

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 Gaussian Process (GP) emulation has become a standard for optimizing systems which evaluations are prohibitively expensive to evaluate, 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 methodological contributions ranging from theoretical to methodological questions (such as parallelization, handling large data sets, etc.) and concrete 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. Figure 19 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 lead to generate objective functions that are to be made 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. This concerns notably the incorporation of mathematical invariances such as symmetries or harmonicity and also links to global sensitivity analysis, paving the way to novel high-dimensional GP emulation approaches [1, 3].

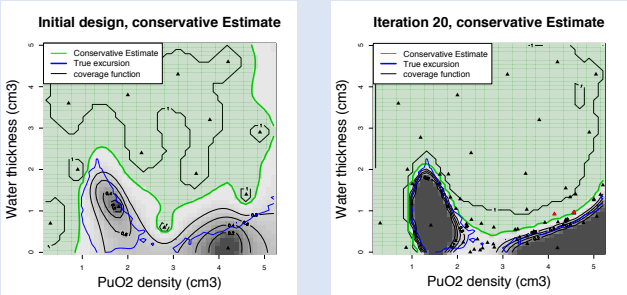


Figure 20

Excursion set estimates obtained on a neutronic criticality safety test case.

COMPUTER EXPERIMENTS FOR THE QUANTIFICATION AND THE REDUCTION OF UNCERTAINTIES

Besides global optimization, the UQOD group has also been focusing on strategies dedicated to other goals such as locating parameter regions leading to a response exceeding a given threshold, that correspond for instance to a dangerous or abnormal behaviour of the considered system. GP-based methods have also been investigated to quantify uncertainties on such sets by using stochastic simulation techniques [2]. Recently, approaches have been proposed for conservative estimation of such sets, i.e. by controlling for instance the probability, under GP assumptions, for set estimates to be completely included in the target set. Figure 20 presents 95% conservative set estimates obtained on a neutronic criticality safety test case (from the French Institut de Radioprotection et de Sûreté Nucléaire) starting from an initial design of experiments (left) and after a novel sequential strategy (right) dedicated to this class of conservative set estimation problem [5]. This ongoing work, in collaboration with colleagues from Neuchâtel, CentraleSupélec and IRSN, was presented at the SIAM 2016 conference on Uncertainty Quantification. For his PhD thesis entitled "Bayesian Set Estimation Relying on Random Field Priors" defended in November 2016 at the University of Bern, Dario Azzimonti received the C. Moser Award for outstanding Ph.D. or M.Sc. theses of IMSV. In other respects, UQOD has been active in designing computer experiments based on distances between set-valued inputs for output distribution quantization [4].

EXTREME VALUE ANALYSIS FOR NON-STATIONARY TIME SERIES WITH APPLICATION IN CLIMATE SCIENCES

Extreme value analysis is a key approach to a number of phenomena from nature, economic and industry activities. In order to assess risks and make reliable decisions, it is of importance to understand and forecast how extreme some given quantities of interest may be or become, be they for instance climatological or economical variables. 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 to change over time. In approaches currently being developed within UQOD in collaboration with the Oeschger Center for Climate Change Research we tackle extreme value analysis of non-stationary climatological times series. We investigate algorithms that, in contrast with broadly used block-wise maxima approaches, do not only rely on yearly extremes but also take into account higher frequency information in order to better capture non-stationarities (See Figure 21).

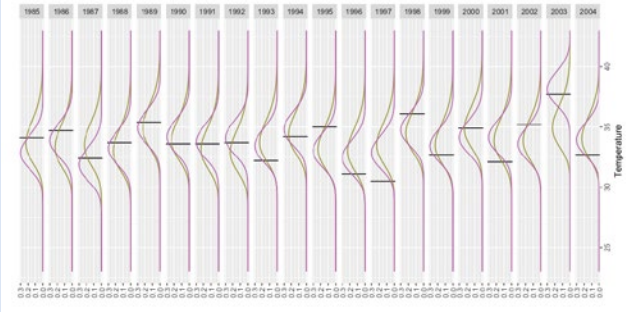


Figure 21

Annual maximum temperatures at Geneva station between 1985 and 2004: observed values (segments) versus retrospectively fitted probability distributions (gray and purple densities).

KEY PUBLICATIONS

[1] On degeneracy and invariances of random fields paths with applications in Gaussian process modelling, David Ginsbourger, Olivier Roustant and Nicolas Durrande, in: Journal of Statistical Planning and Inference, 170:117–128, 2016.

[2] Quantifying uncertainties on excursion sets under a Gaussian random field prior, Dario Azzimonti, Julien Bect, Clément Chevalier and David Ginsbourger, in: SIAM/ASA J. Uncertainty Quantification, 4(1):850- 874, 2016.

[3] On ANOVA Decompositions of Kernels and Gaussian Random Field Paths, David Ginsbourger, Olivier Roustant, Dominic Schuhmacher, Nicolas Durrande and Nicolas Lenz, in: Monte Carlo and Quasi-Monte Carlo Methods, pages 315–330, Springer International Publishing, 2016.

[4] Design of Computer Experiments Using Competing Distances Between Set-Valued Inputs, David Ginsbourger, Jean Baccou, Clément Chevalier and Frédéric Perales, in: mODa 11 – Advances in Model- Oriented Design and Analysis, pages 123–131, Springer International Publishing, 2016.

[5] Contributions to Bayesian set estimation relying on random field priors, Dario Filippo Azzimonti, Inauguraldissertation der Philosophisch-naturwissenschaftlichen Fakultät der Universität Bern, 2016.

7. Computational Bioimaging

OVERVIEW

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

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 2016, the Computational Bioimaging Group was composed of the head of group, three PhD students and one intern.

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.

DEVELOPMENT OF OPTICAL PROJECTION TOMOGRAPHY ALGORITHMS TO ENABLE USE OF HIGH-RESOLUTION MICROSCOPY OBJECTIVES

Traditional optical projection tomography requires the use of low numerical aperture objectives in order to achieve a large depth of field and approximate parallel projection geometry. However, low NA objectives suffer from poor resolution, resulting in blur in the reconstructed image. We developed a modified filtered backprojection method suitable for

focal-plane-scanning optical projection tomography (FPS-OPT), where each projection is obtained by scanning through focal planes during collection. We showed that FPS-OPT has an exact inversion formula akin to a filtered backprojection (FBP), but that incorporates the system's point-spread-function to recover a deblurred 3D volume. With simulations, we demonstrated that FPS-OPT permits the use of high numerical aperture objectives that lead to more accurate images. We further illustrated the technique on experimentally acquired data from a fluorescently-labeled zebrafish larva, which shows that our approach reduces out-of-focus blur.

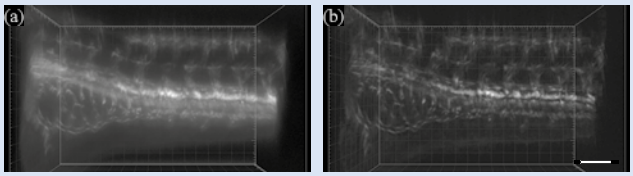


Figure 22

We compare (a) standard FBP with (b) our PSF-aware FBP for 3D fluorescence OPT imaging of a Tg(fli1a:eGFP) zebrafish tail. Scalebar is 100 m. Adapted from [1].

SIMULTANEOUS TEMPORAL SUPERRESOLUTION AND DENOISING FOR CARDIAC FLUORESCENCE MICROSCOPY

Due to low light emission of fluorescent samples, live fluorescence microscopy imposes a tradeoff between spatiotemporal resolution and signal-to-noise ratio. This can result in images and videos containing motion blur or Poisson-type shot noise, depending on the settings used during acquisition. We proposed an algorithm to simultaneously denoise and temporally super-resolve movies of repeating microscopic processes that is compatible with any conventional microscopy setup that can achieve imaging at a rate of at least twice that of the fundamental frequency of the process (above 4 frames per second for a 2 Hz process). Our method combines low temporal resolution frames from multiple cycles of a repeating process to reconstruct a denoised, higher temporal resolution image sequence which is the solution to a linear

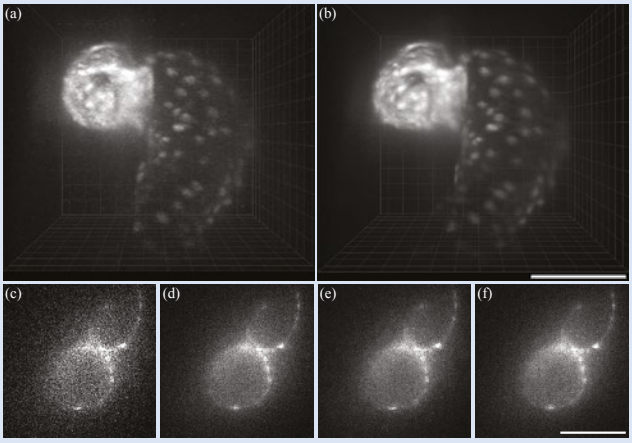


Figure 23

We imaged the heart of a live, 2.5 day old zebrafish embryo imaged in fluorescence at 60 frames per second. (a) low illumination intensity and short integration time used during acquisition result in an image severely corrupted by Poisson-type noise. (b) Our temporal superresolution reconstruction is able to simultaneously temporally superresolve the image sequence and remove much of the noise. Scale bar is 100 m. Adapted from [2].

program that maximizes the consistency of the reconstruction with the measurements, under a regularization constraint. We described, in particular, a parallelizable superresolution reconstruction algorithm and demonstrated its application to live cardiac fluorescence microscopy. Using our method, we experimentally showed temporal resolution improvement by a factor of 1.6, resulting in a visible reduction of motion blur in both on-sample and off-sample frames.

KEY PUBLICATIONS

[1] Kevin G. Chan and Michael Liebling, " A Point-Spread-Function-Aware Filtered Backprojection Algorithm for Focal-Plane-Scanning Optical Projection Tomography, " 2016 IEEE International Symposium on Biomedical Imaging, 2016.

[2] Kevin G. Chan, Sebastian J. Streichan, Le A. Trinh and Michael Liebling " Simultaneous temporal superresolution and denoising for cardiac fluorescence microscopy, " IEEE Transactions on Computational Imaging, 2016.

8. Biometric Person Recognition

OVERVIEW

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

GROUP OVERVIEW

The Biometrics group investigates and develops novel image-processing and patternrecognition algorithms for face recognition (2D, 3D, and near-infrared), speaker recognition, vein recognition, anti-spoofing (PAD: presentation attack detection), and emerging biometric modes (EEG and remote photoplethysmography). The group is geared toward reproducible research and technology transfer, using its own signal-processing and machine learning toolbox Bob (www.idiap.ch/software/bob/).

The Biometrics group in 2016 was composed of 1 head of group, 3 research associates, 3 postdocs, 3 PhD students, 1 intern and 1 visitor.

KEY SCIENTIFIC OUTPUTS

The Biometrics 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 biometrics 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.

HETEROGENEOUS FACE RECOGNITION

The task of Heterogeneous Face Recognition (Figure 24) 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.

Figure 24

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).

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 25). 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 26) 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 25

Illustration of video and audio presentation attacks.

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 27).

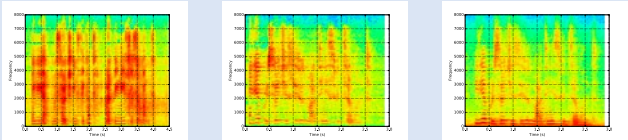


Figure 26

Spectrograms of genuine samples (left) vs. attack (middle and right) samples.

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,

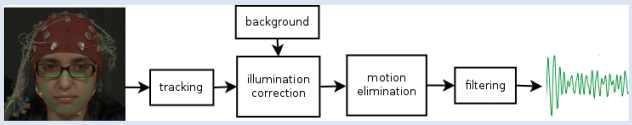


Figure 27

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

wavelength between 700 and 1000 nm) illumination. The vein pattern can then be extracted with image pre-processing techniques (Figure 28) and used for recognition by any pattern recognition method.

SWISS CENTER FOR BIOMETRICS RESEARCH AND TESTING

In 2014, the Idiap Research Institute launched the " Swiss Center for Biometrics Research and Testing " (www.biometrics-center.ch), a competence center within the Institute following recent successes in coordinating International research projects in Biometrics (MOBIO, TABULA RASA and BEAT). The aim of this center is to serve as a legacy for these projects and to push for industry-driven research and testing in biometrics. The center attracted the attention of large companies (license, research and testing agreements) and led to many new projects (DARPA, IARPA, CTI).

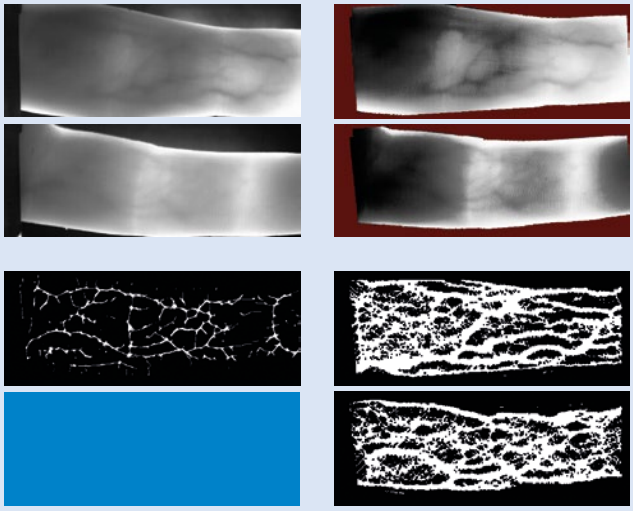


Figure 28

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).

KEY PUBLICATIONS

[1] P. Korshunov and S. Marcel, " Cross-database evaluation of audio-based spoofing detection systems ", Proceedings of Interspeech, 1705–1709, 2016.

[2] M. Günther and L. El Shafey and S. Marcel, " Face Recognition in Challenging Environments: An Experimental and Reproducible Research Survey ", Face Recognition Across the Imaging Spectrum, T. Bourlai (Eds), pages 247-280, 2016.

[3] I. Chingovska and N. Erdogmus and A. Anjos and S. Marcel, " Face Recognition Systems Under Spoofing Attacks ", Face Recognition Across the Imaging Spectrum, T. Bourlai (Eds), pages 165–194, 2016.

[4] A. Hadid, N. Evans, S. Marcel, and J. Fierrez, " Biometrics systems under spoofing attack: an evaluation methodology and lessons learned ", IEEE Signal Processing Magazine, 32(5):20–30, 2016.

[5] A. Morales, J. Fierrez, R. Tolosana, J. Ortega-Garcia, J. Galbally, M. Gomez-Barrero,A. Anjos and S. Marcel, " Keystroke Biometrics Ongoing Competition ", IEEE Access, Vol 4, 7736–7746, 2016.

9. 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 2016, the NLP group had the following members: the head of the group, three postdoctoral students, and two PhD students. A PhD thesis was defended at EPFL in February 2016.

KEY SCIENTIFIC OUTPUTS

In 2016, we have demonstrated that the semantic analysis of noun phrases and pronouns in documents with multiple sentences are beneficial to statistical MT, an innovation that has been met with interest from the community, following our earlier work on discourse connectives and verb tenses. We have shown that a multiple-instance learning framework was appropriate to predict aspect ratings in texts (e.g. sentiments towards aspects of items), and was able to explain the predictions through the weights given to sentences, which can be used for segmentation and summarization as well. Recently, we designed hierarchical neural networks with attention, enabling cross-lingual transfer for document classification in 8 languages.

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

DOCUMENT-LEVEL MACHINE TRANSLATION

We focus on a problem that is not often targeted in the current statistical machine translation (SMT) paradigm: the translation of words that depend on long-range relationships between sentences. The main examples are discourse connectives (which have been our main focus in 2010–2014)[23], verb tenses (studied in 2013–2014), and pronouns and noun phrases (our current focus)[24]. The NLP group coordinates a consortium with teams from Geneva, Zürich and Utrecht. In 2016, our main achievements in document-level noun and pronoun MT were the following ones.

We designed a method to enforce the consistency of noun translation, as illustrated in Fig. 29. The method is based on a classifier that decides whether two occurrences of the same noun in a source text should be translated consistently. We combined these classifiers with phrase-based SMT for Chineseto- English and German-to-English, in two ways: automatic post-editing of noun translations vs. reranking 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. First, we built a coreference model [1], which captures the probabilistic connection between a pronoun and the features of its antecedent (gender, number and humanness) by learning from the output

[23] Supported by the COMTIS SNSF Sinergia project, 2010-2013, see www.idiap.ch/en/scientific-research/projects/COMTIS

[24] Supported by the MODERN SNSF Sinergia project, 2013-2017, see www.idiap.ch/project/modern.

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. Moreover, in a knowledge-poor approach, we built a pronoun-aware language model that is used to re-rank MT hypotheses [2]. The model estimates the probability of a target pronoun given the gender and number of preceding nouns or pronouns on the target side only. This improved the accuracy of pronoun translation by 5% over the baseline. In the above work, to evaluate the translation of pronouns, we compared them automatically with those in a reference translation by using our APT metric (Accuracy of Pronoun Translation). Using results from the DiscoMT 2015 shared task, we showed that APT correlates well with human judgments.

Example 1
Source: nach einführung 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: 欺诈性旅行或身份证件系指有下列情形之一的任何旅行或身份证件
Reference: Fraudulent travel or identity **document**; shall mean any travel or identity **document**
MT: 欺诈性 travel or identity **papers**. 系指 have under one condition; any travel, or identity **document**

Figure 29

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.

Finally, 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 mentiongrouping 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.

TEXT ANALYSIS FOR MULTIMEDIA RECOMMENDATION

We develop methods for text and multimedia information retrieval and recommendation, based on the analysis of content and sentiment, and on the networked structure of (social) media repositories. This work was supported by the SUMMA EU project[25], TheArk Foundation[26], the CTI/KTI[27], and included a collaboration with the EPFL Social Media Lab.

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 [3], we demonstrated the explanatory power of our model by comparing its predicted aspect saliency values with those assigned by humans. To this end, we collected human aspect values over a set of audiobook reviews, and made them available as the HATDOC dataset [4]. Moreover, 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,

[25] See www.summa-project.eu.

[26] Technology transfer project with Unono net, a university social network (www.unono.net).

[27] Technology transfer project with Faveeo SA, an information watch company (www.faveeo.com).

in 8 languages, with multiple topic labels. We thus delivered a deep multilingual tagging library for the SUMMA demonstrator, illustrated in Fig. 30 for one test article.

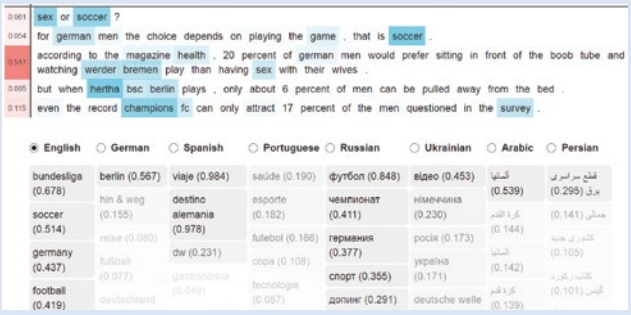


Figure 30

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

Working on personalized search for information watch with Faveeo SA, we designed solutions for two recommendation tasks. First, we designed a method to suggest additional search terms to users, related to their initial query terms but diversifying them too. Our comparative study of two word representations, probabilistic topic models vs. word embeddings, showed the superiority of the latter [5]. Moreover, we designed a document recommender, which extends the initial set of results from Faveeo based on what users 'liked', using low-dimensional vector representations of documents.

Within two focused collaborations, we applied our know-how of text similarity methods to two problems. We designed for Unono.net modules for recommending news items to users of their portal, and for recommending profiles (i.e. CVs of applicants) to employers based on their job descriptions, and viceversa. In collaboration with the EPFL Social Media Lab, we designed a module for extracting and clustering n-grams from social media, and visualizing the results. Exploring results under a range of parameters, our tools helped to observe the propagation of opinions about 6 French TV series in social media.

KEY PUBLICATIONS

[1] Luong N.Q. & Popescu-Belis A. (2016) – Improving Pronoun Translation by Modeling Coreference Uncertainty. Proceedings of WMT 2016 (First Conference on Machine Translation), Research Papers, Berlin, Germany, p. 12–20.

[2] Luong N.Q. & Popescu-Belis A. (2016) – A Contextual Language Model to Improve Machine Translation of Pronouns by Re-ranking Translation Hypotheses. Proceedings of EAMT 2016 (19th Annual Conference of the European Association for Machine Translation), Riga, Latvia, p. 292–304.

[3] Pappas N. & Popescu-Belis A. (2016) – Human versus Machine Attention in Document Classification: A Dataset with Crowdsourced Annotations. Proceedings of the EMNLP SocialNLP workshop (4th Int. Workshop on Natural Language Processing for Social Media), Austin, TX, p. 94–100.

[4] Pappas N. & Popescu-Belis A. (2016) – Adaptive Sentiment-Aware One-Class Collaborative Filtering. Expert Systems With Applications, vol. 43, p. 23–41.

[5] Habibi M., Mahdabi P. & Popescu-Belis A. (2016) – Question Answering in Conversations: Query Refinement Using Contextual and Semantic Information. Data & Knowledge Engineering, vol. 106, p. 38–51.

Projects in progress during 2016

PROJECTS FUNDED BY EUROPEAN AND INTERNATIONAL AGENCIES

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

---

Name **BEAT** (Biometrics Evaluation and Testing)
Funding FP7 – STREP – SEC
Coordinator Idiap Research Institute
Duration 2012.03.01 – 2016.02.28
Partner(s) École Polytechnique Fédérale de Lausanne, University of Surrey, Commissariat à l'Énergie Atomique, Universidad Autónoma de Madrid, Katholieke Universiteit Leuven, Sagem Securité SA, TUBITAK Informatics and Information Security, Advanced Technologies Research Center, TÜViT

Name **EUMSSI** (Event Understanding through Multimodal Social Stream Interpretation)
Funding FP7 – IP – ICT
Coordinator Fundacio Barcelona Media
Duration 2013.12.01 – 2016.11.30
Partner(s) Idiap Research Institute, Université du Maine, Gottfried Wilhelm Leibniz Universität Hannover, Gesellschaft zur Förderung der angewandten Informationsforschung EV, Deutsche Welle, Video Stream Networks SL

Name **SIIP** (Speaker Identification Integrated Project)
Funding FP7 – SEC – 2013 – 1
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 OfWarwick, Laboratorio Di Scienze Della Cittadinanza, The International Criminal Police Organization, Police Service Of Northern Ireland, Ministério Da Justiça, Lisboa

---

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

Name	**MUMMER** (MultiModal Mall Entertainment Robot)
Funding	H2020 – RIA – ICT
Coordinator	University of Glasgow
Duration	2016.03.01 – 2020.02.28
Partner(s)	Idiap Research Institute, Centre national de la recherche scientifique, Aldebarab Robotics, Teknologian Tutkimuskekus VTT, Kiinteistö Oy Ideapark AB

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

Name	**4DHEART** (4D analysis of heart development and regeneration using advanced light microscopy)
Funding	H2020 – MSCA – ITN
Coordinator	Fundacion Centro Nacional de Investigaciones Cardiovasculares Carlos III
Duration	2016.10.01 – 2020.09.30
Partner(s)	Idiap Research Institute, Centre national de la recherche scientifique, Universität Bern, Acquifer AG, Bitplane AG, Leica Microsystems CMS GmbH, 4D-Nature Imaging Consulting, S. L., Centre Européen de Recherche en Biologie et Médecine

Name	**DEXROV** (Effective Dexterous ROV Operations in Presence of Communications Latencies)
Funding	H2020 – RIA – 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	**MALORCA** (Machine Learning of Speech Recognition Models for Controller Assistance)
Funding	H2020 – SESAR – RIA
Coordinator	Deutsches Zentrum für Luft- und Raumfahrt EV
Duration	2016.04.01 – 2018.03.31
Partner(s)	Idiap Research Institute, Universität des Saarlandes, Rízení letového provozu Ceské republiky, Austro Control Österreichische Gesellschaft für Zivilluftfahrt MBH

Name	**SWAN** (Secure Access Control over Wide Area Network)
Funding	Research Council of Norway ICT initiative
Coordinator	Hogskolen I Gjovik
Duration	2016.01.01 – 2018.12.31
Partner(s)	Idiap Research Institute, MORPHO, Bankenverband, Universitetet Oslo, Zwipe AS

Name	**SAVI** (Media Forensics)
Funding	DARPA
Coordinator	SRI International
Duration	2016.05.19 – 2020.05.18
Partner(s)	Idiap Research Institute

PROJECTS FUNDED BY SWISS AGENCIES

Name **25E** (Célébrations du 25e anniversaire de l'Idiap)
Funding Loterie Romande
Coordinator Idiap Research Institute
Duration 2016.09.01 – 2016.09.30

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 **LIVEHEART** (The Cellular Basis of Cardiac Development Revealead by Live Imaging)
Funding SNSF – ANR
Coordinator Institut de génétique et de biologie moléculaire et cellulaire
Duration 2016.03.01 – 2019.02.28
Partner(s) Idiap Research Institute, Universität Bern, École Polytechnique Paris

---

Name **ODESSA** (Online Diarization Enhanced by recent Speaker identification and Sequential learning Approaches)
Funding SNSF – ANR
Coordinator Centre national de la recherche scientifique
Duration 2016.03.01 – 2019.02.28
Partner(s) Idiap Research Institute, EURECOM

Name **TACT-HAND** (Improving control of prosthetic hands using tactile sensors and realistic machine learning)
Funding SNSF – DACH
Coordinator Deutsches Zentrum für Luft- und Raumfahrt EV
Duration 2016.04.01 – 2019.03.31
Partner(s) Idiap Research Institute, Universität Bielfeld

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

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

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

Name **HFACE** (Heterogeneous Face Recognition)
Funding SNSF – Division II
Coordinator Idiap Research Institute
Duration 2014.07.01 – 2018.06.30

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

Name **A-MUSE** (Adaptive Multilingual Speech Processing)
Funding SNSF – Division II
Coordinator Idiap Research Institute
Duration 2012.10.01 – 2016.09.30

Name **DASH** (Object Detection with Active Sample Harvesting)
Funding SNSF – Division II
Coordinator Idiap Research Institute
Duration 2012.10.01 – 2016.09.30

Name **PHASER** (Parsimonious Hierarchical Automatic Speech Recognition)
Funding SNSF – Division II
Coordinator Idiap Research Institute
Duration 2014.06.01 – 2016.05.31

Name **I-DRESS** (Assistive Interactive robotic system for support in DRESSing)
Funding SNSF – ERA-NET Chist Era
Coordinator Institut de Robotica i Informatica Industrial, CSIC-UPC
Duration 2015.12.01 – 2018.11.30
Partner(s) Idiap Research Institute, University of the West of England

---

Name **YOUTH@NIGHT** (A multi-disciplinary multi-method study of young people's outgoing and drinking behaviors)
Funding SNSF – Interdisciplinary project
Coordinator Sucht Schweiz – Research Institute
Duration 2014.01.01 – 2017.12.31
Partner(s) Idiap Research Institute, Universität Zürich

Name **PLATFORM-MMD** (Platform for Reproducible Acquisition, Processing, and Sharing of Dynamic, Multi-Modal Data)
Funding SNSF – R'equip
Coordinator Idiap Research Institute
Duration 2016.07.01 – 2017.06.30

Name **SP2** (Speech Prosody)
Funding SNSF – Scopes
Coordinator Idiap Research Institute
Duration 2014.04.01 – 2016.05.31
Partner(s) Budapest University of Technology and Econom, Ss. Cyril and Methodius University, Universizy of Novi Sad

---

Name **SMILE** (Scalable Multimodal sign language Technology for sIgn language Learning and assessmEnt)
Funding SNSF – 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

---

Name **SIWIS** (Spoken Interaction with Interpretation in Switzerland)
Funding SNSF – Sinergia
Coordinator Idiap Research Institute
Duration 2012.12.01 – 2016.11.30
Partner(s) University of Edinburgh, University of Geneva, Eidgenössische Technische Hochschule Zürich

---

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

Name **UBIMPRESSED** (Ubiquitous First Impressions and Ubiquitous Awareness)
Funding SNSF – Sinergia
Coordinator Idiap Research Institute
Duration 2014.01.01 – 2017.06.30
Partner(s) Université de Neuchâtel, Cornell University

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

---

Name **OMSI-EXT** (Objective Measurement of Speech Intelligibility)
Funding ArmaSuisse
Coordinator Idiap Research Institute
Duration 2016.04.01 – 2016.12.31

Name **MIRROR** (Mobile Data to Improve Decision Making in African Cities)
Funding EPFL collaboration
Coordinator Idiap Research Institute
Duration 2014.01.01 – 2016.12.31
Partner(s) École Polytechnique Fédérale de Lausanne

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 énergétiques et municipales
Duration 2016.11.01 – 2017.03.31
Partner(s) Idiap Research Institute

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 **L-PASS** (Linguistic-Paralinguistic Speech Synthesis)
Funding Hasler Foundation
Coordinator Idiap Research Institute
Duration 2015.12.01 – 2016.02.28

Name **RECAPP** (Making speech technology accessible to Swiss people)
Funding Hasler Foundation
Coordinator Idiap Research Institute
Duration 2015.01.01 – 2016.12.31

Name **MASH-2** (Massive Sets of Heuristics for Machine Learning II)
Funding Hasler Foundation
Coordinator Idiap Research Institute
Duration 2013.07.01 – 2017.06.30

Name **ADDG2SU** (Flexible Acoustic Data-Driven Grapheme to Subword Unit Conversion)
Funding Hasler Foundation
Coordinator Idiap Research Institute
Duration 2013.03.01 – 2017.02.28

Name **SODS** (Semantically Self-Organized Distributed Web Search)
Funding Hasler Foundation
Coordinator Idiap Research Institute
Duration 2012.01.01 – 2016.04.30

PROJECTS FUNDED BY INDUSTRIAL PARTNERS

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

Name **BIOWAVE** (A BIOmetric Watch Activated by VEins)
Funding CTI
Coordinator Centre Suisse d'Electronique et de Microtechnique
Duration 2016.02.01 – 2017.07.31
Partner(s) Idiap Research Institute, Biowatch 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 **FARGO** (Convenient and Secure 3D Face Recognition based on RGB-D Cameras)
Funding CTI
Coordinator Idiap Research Institute
Duration 2016.05.01 – 2017.10.31
Partner(s) KeyLemon SA

Name **IMIM** (Intelligent Monitoring for In-line Manufacturing)
Funding CTI
Coordinator AISA Automation Industrielle SA
Duration 2016.07.01 – 2018.02.28
Partner(s) Idiap Research Institute

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 **MOGS** (2e génération de l'app MOGS)
Funding CTI
Coordinator Mogs SA
Duration 2016.03.01 – 2016.05.31
Partner(s) Idiap Research Institute

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 2015.12.01 – 2017.07.31
Partner(s) DomoSafety S.A., Universität Bern, HES-SO Vaud

Name **SCOREL2** (Automatic scoring and adaptive pedagogy for oral language learning)
Funding CTI
Coordinator Idiap Research Institute
Duration 2014.06.01 – 2016.01.31
Partner(s) SpeedLingua

Name **FAVEO** (Accelerating online information discovery through context-driven and behaviour-based personalization of search)
Funding CTI
Coordinator Idiap Research Institute
Duration 2014.09.01 – 2016.03.31
Partner(s) École Polytechnique Fédérale de Lausanne, Faveeo

Name **VIDEOPROTECTOR** (Morphean VideoProtector)
Funding CTI
Coordinator Idiap Research Institute
Duration 2014.06.01 – 2016.01.31
Partner(s) HES-SO Fribourg

Name **DIGIT-ARENA** (Real-Time Perimeter Board Content Digital Replacement)
Funding CTI
Coordinator Idiap Research Institute
Duration 2014.12.01 – 2016.07.30
Partner(s) Digit Arena

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** (Patrouille des Glaciers – fatigue)
Funding Fondation The Ark
Coordinator Clinique romande de réadaptation
Duration 2016.08.01 – 2017.03.31
Partner(s) Idiap Research Institute, Life

Name **AUDIOSEARCHDEMO** (Demonstration for AudioSearch)
Funding Fondation The Ark
Coordinator Idiap Research Institute
Duration 2016.05.01 – 2016.07.31

Name **EYEWARE** (Eyeware)
Funding Fondation The Ark
Coordinator Idiap Research Institute
Duration 2015.09.01 – 2016.04.30

Name **HYBRID** (Hybrid Recommender System)
Funding Fondation The Ark
Coordinator Unono
Duration 2015.08.01 – 2016.03.31
Partner(s) Idiap Research Institute

Name **PUNK** (Punktuation)
Funding Fondation The Ark
Coordinator recapp
Duration 2015.05.01 – 2016.04.30
Partner(s) Idiap Research Institute

Name **SUVA** (Recomed, Intégration de la transcription vocale dans le dossier patient informatisé CRR)
Funding Fondation The Ark
Coordinator Idiap Research Institute
Duration 2014.03.01 – 2016.01.31
Partner(s) Cimark, Clinique romande de réadaptation

Name **NMTBENCHMARK** (Training and Benchmarking Neural MT and ASR Systems for Swiss Languages)
Funding Industrial Project
Coordinator Idiap Research Institute
Duration 2016.12.01 – 2017.11.30
Partner(s) École Polytechnique Fédérale de Lausanne

Name **SNACK** (Understanding how to balance the enjoyment of food with the effects of calories)
Funding Industrial Project
Coordinator École Polytechnique Fédérale de Lausanne
Duration 2016.03.01 – 2018.02.28
Partner(s) Idiap Research Institute

Name **DRACULA** (Detect and track people/object in order to deliver personalised movies.)
Funding Industrial Project
Coordinator Idiap Research Institute
Duration 2015.07.01 – 2016.02.29
Partner(s) iRewind

Name **GOOGLE-MOBILE** (Mobile Face and Voice Anti Spoofing)
Funding Industrial Project
Coordinator Idiap Research Institute
Duration 2015.05.15 – 2016.03.31

PROJECTS AWARDED IN 2016 AND STARTING IN 2017

Name **SIMPLICID** (Seemless Customer Identification)
Funding CTI
Coordinator Idiap Research Institute
Duration 2017.03.01 – 2018.11.30
Partner(s) Swisscard AECS GmbH

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 **COBALT** (Content Based Call Filtering)
Funding Fondation The Ark
Coordinator Idiap Research Institute
Duration 2017.03.01 – 2018.02.28
Partner(s) Katia SA

Name **RECOMEDIND** (RecoMed Industrialisation)
Funding Fondation The Ark
Coordinator recapp
Duration 2017.03.01 – 2017.07.31
Partner(s) Idiap Research Institute, Clinique romande de réadaptation

---

Name **FLOSS** (Flexible Linguistically-guided Objective Speech aSessment)
Funding Hasler Foundation
Coordinator Idiap Research Institute
Duration 2017.03.01 – 2020.02.2

Name **CCSUMMARY** (Classification and Summarization of Written and Spoken Interactions with Customers)
Funding IND – Industrial Project
Coordinator Idiap Research Institute
Duration 2017.03.01 – 2020.02.29

Name **ISUL** (Importance sampling for large-scale unsupervised learning)
Funding SNSF Division II
Coordinator Idiap Research Institute
Duration 2017.03.01 – 2019.02.28

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

Books and Book Chapters

[1] O. Aran, I. Ari, A. Kindiroglu, P. Santemiz, and L. Akarun, "Otomatik i¸saret dili tanıma ve türk i¸saret dili için bilgisayar uygulamaları," in Ellerle Konusmak: Türk ˙I¸saret Dili Ara¸stırmaları / Research on Turkish Sign Language, in Turkish, Koc University Press, 2016, pp. 471–498.

[2] A. G. Billard, S. Calinon, and R. Dillmann, "Learning from humans," in Handbook of Robotics, B. Siciliano and O. Khatib, Eds., 2nd Edition, Secaucus, NJ, USA: Springer, 2016, ch. 74, pp. 1995– 2014.

[3] I. Chingovska, N. Erdogmus, A. Anjos, and S. Marcel, "Face recognition systems under spoofing attacks," in Face Recognition Systems Under Spoofing Attacks, 1st, Springer International Publishing, Feb. 2016, ch. 8, pp. 165–194.

[4] D. Gatica-Perez, S. Ruiz-Correa, and D. Santani, "What tripadvisor can't tell: Crowdsourcing urban impressions for whole cities," in Digital Polis, A. d. Biase, N. Ottaviano, and O. Zaza, Eds., L'Oeil d'Or (translated to French.), 2016.

[5] D. Ginsbourger, J. Baccou, C. Chevalier, and F. Perales, "Design of computer experiments using competing distances between set-valued inputs," in MODa 11 - Advances in Model-Oriented Design and Analysis, ser. Contributions to Statistics, Springer International Publishing, 2016, pp. 123–131.

[6] D. Ginsbourger, O. Roustant, D. Schuhmacher, N. Durrande, and N. Lenz, "On anova decompositions of kernels and gaussian random field paths," in Monte Carlo and Quasi-Monte Carlo Methods, ser. Springer Proceedings in Mathematics & Statistics, vol. 163, Springer International Publishing, 2016, pp. 315–330.

[7] M. Günther, L. El Shafey, and S. Marcel, "Face recognition in challenging environments: An experimental and reproducible research survey," in Face Recognition Across the Imaging Spectrum, T. Bourlai, Ed., 1st ed., Springer, Feb. 2016.

Articles in Scientific Journals

[1] A. Asaei, H. Bourlard, M. J. Taghizadeh, and V. Cevher, "Computational methods for underdetermined convolutive speech localization and separation via model-based sparse component analysis," Speech Communication, vol. 76, pp. 201–217, 2016.

[2] A. Asaei, M. J. Taghizadeh, S. Haghighatshoar, B. Raj, H. Bourlard, and V. Cevher, "Binary sparse coding of convolutive mixtures for sound localization and separation via spatialization," IEEE Transactions on Signal Processing, vol. 64, no. 3, pp. 567–579, 2016.

[3] U. Avci and O. Aran, "Predicting the performance in decision-making tasks: From individual cues to group interaction," IEEE Transactions on Multimedia, vol. 18, no. 4, pp. 643–658, 2016.

[4] D. Azzimonti, J. Bect, C. Chevalier, and D. Ginsbourger, "Quantifying uncertainties on excursion sets under a gaussian random field prior," SIAM/ASA J. Uncertainty Quantification, vol. 4, no. 1, pp. 850–874, 2016.

[5] S. Calinon, "A tutorial on task-parameterized movement learning and retrieval," Intelligent Service Robotics, vol. 9, no. 1, pp. 1–29, Jan. 2016.

[6] G. Can, J.-M. Odobez, and D. Gatica-Perez, "Evaluating shape representations for maya glyph classification," ACM Journal on Computing and Cultural Heritage (JOCCH), vol. 9, no. 3, Sep. 2016.

[7] A. Cerekovic, O. Aran, and D. Gatica-Perez, "Rapport with virtual agents: What do human social cues and personality explain?" IEEE Transactions on Affective Computing, 2016.

[8] M. Cernak, A. Asaei, and H. Bourlard, "On structured sparsity of phonological posteriors for linguistic parsing," Speech Communication, vol. 84, pp. 36–45, 2016.

[9] M. Cernak, Š. Beňuš, and A. Lazaridis, "Speech vocoding for laboratory phonology," Computer Speech and Language, L. ten Bosch, Ed., 2016.

[10] M. Cernak, A. Lazaridis, A. Asaei, and P. N. Garner, "Composition of deep and spiking neural networks for very low bit rate speech coding," IEEE/ACM Trans. on Audio, Speech and Language Processing, 2016.

[11] K. G. Chan, S. J. Streichan, L. A. Trinh, and M. Liebling, "Simultaneous temporal superresolution and denoising for cardiac fluorescence microscopy," IEEE Transactions on Computational Imaging, 2016, in press.

[12] P. Dighe, A. Asaei, and H. Bourlard, "Sparse modeling of neural network posterior probabilities for exemplar-based speech recognition," Speech Communication: Special Issue on Advances in Sparse Modeling and Low-rank Modeling for Speech Processing, vol. 76, 230–244, Feb. 2016.

[13] M. Ferras, S. Madikeri, and H. Bourlard, "Speaker diarization and linking of meeting data," IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 24, no. 11, pp. 1935–1945, Nov. 2016.

[14] M. Ferras, S. Madikeri, S. Dey, P. Motlicek, and H. Bourlard, "A large-scale open-source acoustic simulator for speaker recognition," IEEE Signal Processing Letters, vol. 23, no. 4, pp. 527 –531, 2016.

[15] M. Focchi, A. del Prete, I. Havoutis, R. Featherstone, D. G. Caldwell, and C. Semini, "High-slope terrain locomotion for torque-controlled quadruped robots," Autonomous Robots, 2016.

[16] K. A. Funes Mora and J.-M. Odobez, "Gaze estimation in the 3d space using rgb-d sensors. towards head-pose and user invariance.," International Journal of Computer Vision, vol. 118, no. 2, pp. 194–216, Jun. 2016, First online: 13 November 2015.

[17] L. Fusco, R. Lefort, K. C. Smith, F. Benmansour, G. Gonzalez, C. Barilari, B. Rinn, F. Fleuret, P. Fua, and O. Pertz, "Computer vision profiling of neurite outgrowth dynamics reveals spatiotemporal modularity of rho gtpase signaling," Journal of Cell Biology, vol. 212, no. 1, pp. 91–111, Jan. 2016.

[18] P. Gay, S. Meignier, P. Deleglise, and J.-M. Odobez, "Crf-based context modeling for person identification in broadcast videos," Frontiers in ICT: Computer Image Analysis, vol. 3, 2016.

[19] D. Ginsbourger, O. Roustant, and N. Durrande, "On degeneracy and invariances of random fields paths with applications in gaussian process modelling," Journal of Statistical Planning and Inference, vol. 170, pp. 117–128, Mar. 2016.

[20] M. Habibi, P. Mahdabi, and A. Popescu-Belis, "Question answering in conversations: Query refinement using contextual and semantic information," Data & Knowledge Engineering Journal, Jun. 2016.

[21] I. Himawan, P. Motlicek, D. Imseng, and S. Sridharan, "Feature mapping using far-field microphones for distant speech recognition," Speech Communication, vol. 83, pp. 1–9, Oct. 2016, A publication of the European Association for Signal Processing (EURASIP) and of the International Speech Communication Association (ISCA).

[22] R. Hu, J.-M. Odobez, and D. Gatica-Perez, "Extracting maya glyphs from degraded ancient documents via image segmentation," Journal on Computing and Cultural Heritage, 2016.

[23] I. Kuzborskij and F. Orabona, "Fast rates by transferring from auxiliary hypotheses," Machine Learning, 2016.

[24] I. Kuzborskij, F. Orabona, and B. Caputo, "Scalable greedy algorithms for transfer learning," Computer Vision and Image Understanding, 2016.

[25] L. Lefakis and F. Fleuret, "Jointly informative feature selection," Journal of Machine Learning Research, 2016.

[26] A. Morales, J. Fierrez, R. Tolosana, J. Ortega-Garcia, J. Galbally, M. Gomez-Barrero, A. Anjos, and S. Marcel, " Keystroke biometrics ongoing competition, " IEEE Access, vol. 4, pp. 7736–7746, Nov. 2016.

[27] L. S. Nguyen and D. Gatica-Perez, " Hirability in the wild: Analysis of online conversational video resumes, " IEEE Trans. on Multimedia, vol. 18, no. 7, pp. 1422–1437, Jul. 2016.

[28] N. Pappas and A. Popescu-Belis, " Adaptive sentiment-aware one-class collaborative filtering, " Expert Systems with Applications, vol. 43, pp. 23–41, Jan. 2016.

[29] R. Rasipuram and M. Magimai.-Doss, " Articulatory feature based continuous speech recognition using probabilistic lexical modeling, " Computer Speech and Language, vol. 36, pp. 233–259, 2016.

[30] M. Razavi, R. Rasipuram, and M. Magimai.-Doss, " Acoustic data-driven grapheme-to-phoneme conversion in the probabilistic lexical modeling framework, " Speech Communication, vol. 80, 2016.

[31] L. Rozo, S. Calinon, D. G. Caldwell, P. Jimenez, and C. Torras, " Learning physical collaborative robot behaviors from human demonstrations, " IEEE Trans. on Robotics, vol. 32, no. 3, pp. 513– 527, Jun. 2016.

[32] L. Rozo, J. Silverio, S. Calinon, and D. G. Caldwell, " Learning controllers for reactive and proactive behaviors in human-robot collaboration, " Frontiers in Robotics and AI, vol. 3, no. 30, pp. 1–11, Jun. 2016.

[33] S. Ruiz-Correa, D. Santani, B. Ramirez Salazar, I. Ruiz-Correa, F. Rendon-Huerta, C. Olmos Carrillo, B. Sandoval Mexicano, A. Arcos Garcia, R. Hasimoto-Beltran, and D. Gatica-Perez, " Sensecityvity: Mobile crowdsourcing, urban awareness, and collective action in mexico, " IEEE Pervasive Computing, 2016.

[34] S. Ruiz-Correa, D. Santani, B. Ramirez Salazar, I. Ruiz-Correa, F. Rendon-Huerta, C. Olmos Carrillo, B. Sandoval Mexicano, A. Arcos Garcia, R. Hasimoto-Beltran, and D. Gatica-Perez, " Sensecityvity: Mobile sensing, urban awareness, and collective action in mexico, " IEEE Pervasive Computing, 2016.

[35] N. Suditu and F. Fleuret, " Adaptive relevance feedback for large-scale image retrieval, " Multimedia Tools and Applications, vol. 75, no. 12, pp. 6777–6807, Jun. 2016.

[36] A. K. Tanwani and S. Calinon, " Learning robot manipulation tasks with task-parameterized semitied hidden semi-markov model, " IEEE Robotics and Automation Letters, vol. 1, no. 1, D. Kulic and D. Lee, Eds., pp. 235–242, Jan. 2016.

[37] R. Ullmann and H. Bourlard, " Predicting the intrusiveness of noise through sparse coding with auditory kernels, " Speech Communication, vol. 76, pp. 186–200, 2016.

[38] J. Velasco, D. Pizarro, J. Macias-Guarasa, and A. Asaei, " Tdoa matrices: Algebraic properties and their application to robust denoising with missing data, " IEEE Transactions on Signal Processing, vol. 64, no. 20, pp. 5242–5254, Jun. 2016.

[39] X.Wang, E. Turetken, F. Fleuret, and P. Fua, " Tracking interacting objects using intertwined flows, " IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016.

[40] 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 Transactions on Pattern Analysis and Machine Intelligence

PhD Theses

[1] R. Lebret, " Building word embeddings for solving natural language processing, " Thèse EPFL, nr 7148, PhD thesis, École Polytechnique Fédérale de Lausanne, Sep. 2016.

[2] J. Legrand, " Word sequence modeling using deep learning: And end-to-end approach and its applications, " PhD thesis, EPFL, Oct. 2016.

[3] D. Palaz, " Towards end-to-end speech recognition, " Thèse EPFL nr 7054, PhD thesis, Ecole polytechnique Fédérale de Lausanne, Jul. 2016.

[4] N. Pappas, " Learning explainable user sentiment and preferences for information filtering, " Thèse EPFL, nr 6920, PhD thesis, École Polytechnique Fédérale de Lausanne, Mar. 2016.

[5] D. Santani, " Computational analysis of urban places using mobile crowdsensing, " These EPFL, nr 7243, PhD thesis, Ecole Polytechnique Federale de Lausanne, Nov. 2016.

[6] R. Ullmann, " Can you hear me now? automatic assessment of background noise intrusiveness and speech intelligibility in telecommunications, " PhD thesis, Sciences et Techniques de l'Ingénieur (STI), Jun. 2016.

Articles in Conference Proceedings

[1] O. Aran, D. Sanchez-Cortes, T.-M.-T. Do, and D. Gatica-Perez, " Anomaly detection in elderly daily behavior in ambient sensing environments, " in Proceedings of the 7th Int. Workshop on Human Behavior Understanding, ACM Multimedia, 2016, Amsterdam, Netherlands, 2016.

[2] A. Asaei, M. Cernak, and M. Laganaro, " Paos markers: Trajectory analysis of selective phonological posteriors for assessment of progressive apraxia of speech, " in Proceeding on the 7th Workshop on Speech and Language Processing for Assistive Technologies (SLPAT), 2016.

[3] A. Asaei, G. Luyet, M. Cernak, and H. Bourlard, " Efficient posterior exemplar search space hashing exploiting class-specific sparsity structures, " in Interspeech, San Francisco, CA, 2016.

[4] P. Baqué, T. Bagautdinov, F. Fleuret, and P. Fua, " Principled parallel mean-field inference for discrete random fields, " in Proceedings of the IEEE international conference on Computer Vision and Pattern Recognition, 2016.

[5] K. Bascol, R. Emonet, E. Fromont, and J.-M. Odobez, " Unsupervised interpretable pattern discovery in time series using autoencoders, " in IAPR Int. Workshops on Structural and Syntactic Pattern Recognition (SSPR), Nov. 2016.

[6] D. Berio, S. Calinon, and F. F. Leymarie, " Learning dynamic graffiti strokes with a compliant robot, " in Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems, Oct. 2016, pp. 3981–3986.

[7] C. A. Bhatt, A. Popescu-Belis, and M. Cooper, " Audiovisual summarization of lectures and meetings using a segment similarity graph, " in Proceedings of the ACM International Conference on Multimedia Retrieval (ICMR), ACM, New York, NY: ACM Press, 2016.

[8] S. Calinon, " Stochastic learning and control in multiple coordinate systems, " in Intl Workshop on Human-Friendly Robotics, Genoa, Italy, 2016, pp. 1–5.

[9] G. Can, J.-M. Odobez, C. Pallan Gayol, and D. Gatica-Perez, " Ancient maya writings as highdimensional data: A visualization approach, " in Digital Humanities (DH), Krakow, Jul. 2016.

[10] O. Canévet and F. Fleuret, " Large scale hard sample mining with monte carlo tree search, " in Proceedings of the IEEE international conference on Computer Vision and Pattern Recognition, 2016.

[11] O. Canévet, C. Jose, and F. Fleuret, " Importance sampling tree for large-scale empirical expectation, " in Proceedings of the International Conference on Machine Learning (ICML), New-York, Jun. 2016.

[12] M. Cernak, A. Asaei, P.-E. Honnet, P. N. Garner, and H. Bourlard, " Sound pattern matching for automatic prosodic event detection, " in Interspeech, San Francisco, USA, Sep. 2016.

[13] M. Cernak and P. N. Garner, " Phonvoc: A phonetic and phonological vocoding toolkit, " in Interspeech, San Francisco, USA, Sep. 2016.

[14] K. G. Chan and M. Liebling, " A point-spread-function-aware filtered backprojection algorithm for focal-plane-scanning optical projection tomography, " in 2016 IEEE International Symposium on Biomedical Imaging, Apr. 2016.

[15] A. Costa-Pazo, S. Bhattacharjee, E. Vazquez-Fernandez, and S. Marcel, " The replay-mobile face presentation-attack database, " in Proceedings of the International Conference on Biometrics Special Interests Group, Sep. 2016.

[16] N. Dawalatabad, S. Madikeri, C. C. Sekhar, and H. A. Murthy, " Two-pass ib based speaker diarization system using meeting-specific ann based features, " in Proceedings of Interspeech 2016, Sep. 2016, pp. 2199–2203.

[17] S. Dey, S. Madikeri, M. Ferras, and P. Motlicek, " Deep neural network based posteriors for text-dependent speaker verification, " in Proceedings of 2016 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2016), Shanghai: IEEE, Mar. 2016, pp. 5050– 5054.

[18] S. Dey, S. Madikeri, and P. Motlicek, " Information theoretic clustering for unsupervised domainadaptation, " in Proceedings of 2016 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2016), Shanghai: IEEE, Mar. 2016, pp. 5580–5584.

[19] P. Dighe, G. Luyet, A. Asaei, and H. Bourlard, " Exploiting low-dimensional structures to enhance dnn based acoustic modeling in speech recognition, " in Proceedings of 2016 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2016), Shanghai: IEEE, Mar. 2016, pp. 5690–5694.

[20] M. Ferras, S. Madikeri, S. Dey, P. Motlicek, and H. Bourlard, " Inter-task system fusion for speaker recognition, " in Proceeedings of the INTERSPEECH, 2016.

[21] M. Ferras, S. Madikeri, P. Motlicek, and H. Bourlard, " System fusion and speaker linking for longitudinal diarization of tv shows, " in Proceedings of 2016 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2016), Shanghai: IEEE, Mar. 2016, pp. 5495– 5499.

[22] A. Finnerty, S. Muralidhar, L. S. Nguyen, F. Pianesi, and D. Gatica-Perez, " Stressful first impressions in job interviews, " in Proceedings of the 18th ACM International Conference on Multimodal Interaction, Nov. 2016, pp. 325–332.

[23] T. de Freitas Pereira and S. Marcel, " Heterogeneous face recognition using inter-session variability modelling, " in IEEE Computer Society Workshop on Biometrics, Las Vegas – USA: IEEE, Jun. 2016.

[24] T. Gabor Csapo, G. Nemeth, M. Cernak, and P. N. Garner, " Modeling unvoiced sounds in statistical parametric speech synthesis with a continuous vocoder, " in Proc. of EUSIPCO, Budapest, Hungary, 2016.

[25] J. Gancet, P. Weiss, G. Antonelli, M. F. Pfingsthorn, S. Calinon, A. Turetta, C. Walen, D. Urbina, S. Govindaraj, P. Letier, X. Martinez, J. Salini, B. Chemisky, G. Indiveri, G. Casalino, P. Di Lillo, E. Simetti, D. De Palma, A. Birk, A. K. Tanwani, I. Havoutis, A. Caffaz, and L. Guilpain, "Dexterous undersea interventions with far distance onshore supervision: The dexrov project," in IFAC Conference on Control Applications in Marine Systems (CAMS), Trondheim, Norway, Sep. 2016, pp. 414–419.

[26] B. Gerazov and P. N. Garner, "An agonist-antagonist pitch production model," in Lecture Notes in Artificial Intelligence: 18th International Conference, SPECOM 2016, vol. 9811, Budapest, Hungary, Aug. 2016, pp. 84–91.

[27] B. Gerazov, A. Gjoreski, A. Melov, P.-E. Honnet, Z. Ivanovski, and P. N. Garner, "Unified prosody model based on atom decomposition for emphasis detection," in Proceedings of ETAI, Struga, Macedonia, Sep. 2016.

[28] J.-P. Goldman, P.-E. Honnet, R. Clark, P. N. Garner, M. Ivanova, A. Lazaridis, H. Liang, T. Macedo, B. Pfister, M. S. Ribeiro, E. Wehrli, and J. Yamagishi, "The siwis database: A multilingual speech database with acted emphasis," in Proceedings of Interspeech, San Francisco, USA, Sep. 2016, pp. 1532–1535.

[29] L. Guillou, C. Hardmeier, P. Nakov, S. Stymne, J. Tiedemann, Y. Versley, M. Cettolo, B. Webber, and A. Popescu-Belis, "Findings of the 2016 wmt shared task on cross-lingual pronoun prediction," in Proceedings of WMT 2016 (First Conference on Machine Translation), Association for Computational Linguistics, Berlin, Germany, 2016, 525–542.

[30] I. Havoutis and S. Calinon, "Learning assistive teleoperation behaviors from demonstration," in Proc. IEEE International Symposium on Safety, Security and Rescue Robotics, Oct. 2016, pp. 258–263.

[31] P.-E. Honnet and P. N. Garner, "Emphasis recreation for tts using intonation atoms," in 9th ISCA Speech Synthesis Workshop, Sunnyvale, CA, USA, Sep. 2016, pp. 14–20.

[32] R. Hu, J.-M. Odobez, and D. Gatica-Perez, "Assessing a shape descriptor for analysis of mesoamerican hieroglyphics: A view towards practice in digital humanities," in Digital Humanities Conference (DH), Krakow, Jul. 2016.

[33] C. Jose and F. Fleuret, "Scalable metric learning via weighted approximate rank component analysis," in ECCV 2016, 2016.

[34] P. Korshunov and S. Marcel, "Cross-database evaluation of audio-based spoofing detection systems," in Interspeech, San Francisco, USA, Sep. 2016.

[35] P. Korshunov and S. Marcel, "Joint operation of voice biometrics and presentation attack detection," in IEEE International Conference on Biometrics: Theory, Applications and Systems, Niagara Falls, Buffalo New York, USA, Sep. 2016.

[36] P. Korshunov, S. Marcel, H. Muckenhirn, A. R. Gonçalves, A. G. S. Mello, R. P. V. Violato, F. O. Simões, M. U. Neto, M. de Assis Angeloni, J. A. Stuchi, H. Dinkel, N. Chen, Y. Qian, D. Paul, G. Saha, and M. Sahidullah, "Overview of btas 2016 speaker anti-spoofing competition," in IEEE International Conference on Biometrics: Theory, Applications and Systems, Niagara Falls, NY, USA, Sep. 2016.

[37] I. Kuzborskij, F. M. Carlucci, and B. Caputo, "When naïve bayes nearest neighbors meet convolutional neural networks," in Computer Vision and Pattern Recognition (CVPR), IEEE Conference on, Jun. 2016.

[38] A. Lazaridis, M. Cernak, and P. N. Garner, "Probabilistic amplitude demodulation features in speech synthesis for improving prosody," in Proceedings of Interspeech, San Francisco, USA, 2016.

[39] A. Lazaridis, M. Cernak, P.-E. Honnet, and P. N. Garner, "Investigating spectral amplitude modulation phase hierarchy features in speech synthesis," in 9th ISCA Speech Synthesis Workshop, Sunnyvale, CA, USA, Sep. 2016.

[40] A. Lazaridis, I. Himawan, P. Motlicek, I. Mporas, and P. N. Garner, "Investigating cross-lingual multi-level adaptive networks: The importance of the correlation of source and target languages," in Proceedings of the International Workshop on Spoken Language Translation, Seattle, WA, USA, Dec. 2016.

[41] N. Le, A. Heili, and J.-M. Odobez, "Long-term time-sensitive costs for crf-based tracking by detection," in 2ndWorkshop on Benchmarking Multi-target Tracking: MOTChallenge 2016, Amsterdam, Oct. 2016.

[42] N. Le, A. Heili, D. Wu, and J.-M. Odobez, "Temporally subsampled detection for accurate and efficient face tracking and diarization," in International Conference on Pattern Recognition, Cancun, Mexico: IEEE, Dec. 2016.

[43] N. Le, S. Meignier, and J.-M. Odobez, "Eumssi team at the mediaeval person discovery challenge 2016," in MediaEval Benchmarking Initiative for Multimedia Evaluation, Hilversum, Netherlands, Oct. 2016.

[44] N. Le and J.-M. Odobez, "Learning multimodal temporal representation for dubbing detection in broadcast media," in ACM Multimedia, Amsterdam: ACM, Oct. 2016.

[45] R. Lebret, "Building word embeddings for solving natural language processing," Thèse EPFL, nr 7148, PhD thesis, École Polytechnique Fédérale de Lausanne, Jul. 2016.

[46] J. Legrand, "Word sequence modeling using deep learning: And end-to-end approach and its applications," PhD thesis, EPFL, Aug. 2016.

[47] J. Legrand, M. Auli, and R. Collobert, "Neural network-based word alignment through score aggregation," in Proceedings of the ACL 1st Conference on Machine Translation, 2016.

[48] J. Legrand and R. Collobert, "Deep neural networks for syntactic parsing of morphologically rich languages," in Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, 2016.

[49] J. Legrand and R. Collobert, "Phrase representations for multiword expressions," in Proceedings of the 12th Workshop on Multiword Expressions, 2016.

[50] J. Liyanapathirana and A. Popescu-Belis, "Using the ted talks to evaluate spoken post-editing of machine translation," in Proceedings of the 10th Language Resources and Evaluation Conference (LREC), Portoroz, Slovenia, 2016.

[51] N.-Q. Luong and A. Popescu-Belis, "A contextual language model to improve machine translation of pronouns by re-ranking translation hypotheses," in European Association for Machine Translation, May 2016.

[52] N.-Q. Luong and A. Popescu-Belis, "Improving pronoun translation by modeling coreference uncertainty," in Proceedings of the First Conference on Machine Translation (WMT16), Berlin, Germany, Aug. 2016.

[53] N.-Q. Luong and A. Popescu-Belis, "Pronoun language model and grammatical heuristics for aiding pronoun prediction," in Proceedings of the First Conference on Machine Translation (WMT16), Berlin, Germany: ACL, Aug. 2016.

[54] G. Luyet, P. Dighe, A. Asaei, and H. Bourlard, "Low-rank representation of nearest neighbor phone posterior probabilities to enhance dnn acoustic modeling," in Interspeech, 2016.

[55] P. Mahdabi and A. Popescu-Belis, "Comparing two strategies for query expansion in a news monitoring system," in Natural Language Processing and Information Systems: 21st International Conference on Applications of Natural Language to Information Systems, NLDB 2016, ser. Lecture Notes in Computer Science, vol. 9612, Springer-Verlag, 2016, pp. 267–275.

[56] C. Mastalli, I. Havoutis, M. Focchi, C. Semini, and D. G. Caldwell, "Hierarchical planning of dynamic movements without scheduled contact sequences," in Proceedings of the IEEE International Conference of Robotics and Automation, 2016.

[57] H. Muckenhirn, M. Magimai.-Doss, and S. Marcel, "Presentation attack detection using long-term spectral statistics for trustworthy speaker verification," in Proceedings of International Conference of the Biometrics Special Interest Group (BIOSIG), Sep. 2016.

[58] S. Muralidhar, J. M. R. Costa, L. S. Nguyen, and D. Gatica-Perez, "Dites-moi: Wearable feedback on conversational behavior," in Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia, Dec. 2016.

[59] S. Muralidhar, L. S. Nguyen, D. Frauendorfer, J.-M. Odobez, M. Schmid Mast, and D. Gatica- Perez, "Training on the job: Behavioral analysis of job interviews in hospitality," in Proceedings of the 18th ACM International Conference on Multimodal Interaction, Nov. 2016, pp. 84–91.

[60] J. Newling and F. Fleuret, "Fast k-means with accurate bounds," in Proceedings of the International Conference on Machine Learning (ICML), New York, 2016.

[61] J. Newling and F. Fleuret, "Nested mini-batch k-means," in Proceedings of NIPS, 2016.

[62] C. Oertel, J. D. Lopes, Y. Yu, K. A. Funes Mora, J. Gustafson, A. Black, and J.-M. Odobez, "Towards building an attentive artificial listener: On the perception of attentiveness in audio-visual feedback tokens," in Proceedings of the 18th ACM International Conference on Multimodal Interaction, Tokyo, Japan: ACM, Nov. 2016, pp. 21–28.

[63] D. Palaz, "Towards end-to-end speech recognition," Thèse EPFL nr 7054, PhD thesis, Ecole polytechnique Fédérale de Lausanne, 2016.

[64] N. Pappas, "Learning explainable user sentiment and preferences for information filtering," Thèse EPFL, nr 6920, PhD thesis, École Polytechnique Fédérale de Lausanne, Mar. 2016.

[65] N. Pappas and A. Popescu-Belis, "Human versus machine attention in document classification: A dataset with crowdsourced annotations," in Proceedings of the EMNLP 2016 Workshop on Natural Language Processing for Social Media, Austin, USA, 2016.

[66] N. Pappas, M. Topkara, M. Redi, B. Jou, T. Chen, H. Liu, and S.-F. Chang, "Multilingual visual sentiment concept matching," in International Conference on Multimedia Retrieval, 2016.

[67] P. H. O. Pinheiro, T.-Y. Lin, R. Collobert, and P. Dollar, "Learning to refine object segments," in Computer Vision – ECCV 2016, ser. Lecture Notes in Computer Science, vol. 9905, Amsterdam: Springer, Sep. 2016, pp. 75–91.

[68] A. Popescu-Belis, "Manual and automatic labeling of discourse connectives for machine translation (keynote paper)," in TextLink: Structuring Discourse in Multilingual Europe (Handbook of the Second Action Conference), Budapest, Hungary, 2016, pp. 16–20.

[69] X. Pu, M. A. Chatti, H. Thues, and U. Schroeder, "Wiki-lda: A mixed-method approach for effective interest mining on twitter data," in Proceedings of CSEDU 2016, Apr. 2016.

[70] D. Ram, A. Asaei, and H. Bourlard, "Subspace detection of dnn posterior probabilities via sparse representation for query by example spoken term detection," in Interspeech, 2016.

[71] R. Rasipuram, M. Cernak, and M. Magimai.-Doss, "Hmm-based non-native accent assessment using posterior features," in Proceedings of Interspeech, San Francisco, USA, Sep. 2016.

[72] M. Razavi and M. Magimai.-Doss, "Improving under-resourced language asr through latent subword unit space discovery," in Proceedings of Interspeech, 2016.

[73] E. Roman-Rangel, G. Can, S. Marchand-Maillet, R. Hu, C. Pallan Gayol, G. Krempel, J. Spotak, J.-M. Odobez, and D. Gatica-Perez, " Transferring neural representations for low-dimensional indexing of maya hieroglyphic art, " in Proc. ECCV Workshop on Computer Vision for Art Analysis, Amsterdam, Oct. 2016.

[74] D. Santani, " Computational analysis of urban places using mobile crowdsensing, " These EPFL, nr 7243, PhD thesis, Ecole Polytechnique Federale de Lausanne, Sep. 2016.

[75] D. Santani, J.-I. Biel, F. Labhart, J. Truong, S. Landolt, E. Kuntsche, and D. Gatica-Perez, " The night is young: Urban crowdsourcing of nightlife patterns, " in Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, ser. UbiComp '16, ACM, Sep. 2016.

[76] D. Santani, R. Hu, and D. Gatica-Perez, " Innerview: Learning place ambiance from social media images, " in Proceedings of the 24th ACM International Conference on Multimedia, ser. MM '16, ACM, Oct. 2016.

[77] M. Secujski, B. Gerazov, T. Gabor Csapo, V. Delic, P. N. Garner, A. Gjoreski, D. Guennec, Z. Ivanovski, A. Melov, G. Nemeth, A. Stojkovi´c, and G. Szaszak, " Design of a speech corpus for research on cross-lingual prosody transfer, " in Lecture Notes in Artificial Intelligence: 18th International Conference, SPECOM 2016, vol. 9811, Budapest, Hungary, Aug. 2016, pp. 199–206.

[78] A. K. Tanwani and S. Calinon, " Online inference in bayesian non-parametric mixture models under small variance asymptotics, " in NIPS workshop on Advances in Approximate Bayesian Inference, Barcelona, Spain, 2016, pp. 1–5.

[79] R. Ullmann, " "can you hear me now?" — automatic assessment of background noise intrusiveness and speech intelligibility in telecommunications, " PhD thesis, Sciences et Techniques de l'Ingénieur (STI), 2016.

[80] S. Zagoruyko, A. Lerer, T.-Y. Lin, P. H. O. Pinheiro, S. Gross, S. Chintala, and P. Dollar, " A multipath network for object detection, " in Proceedings of the British Machine Vision Conference, BMVA Press, Sep. 2016.

[81] M. Zeestraten, S. Calinon, and D. G. Caldwell, " Variable duration movement encoding with minimal intervention control, " in Proc. of the IEEE Intl Conf. on Robotics and Automation (ICRA), May 2016, pp. 497–503.

[82] M. Zeestraten, A. Pereira, M. Althoff, and S. Calinon, " Online motion synthesis with minimal intervention control and formal safety guarantees, " in Proc. IEEE Intl Conf. on Systems, Man, and Cybernetics, Budapest, Hungary, Oct. 2016.