



RAPPORT ANNUEL 2015



réalisation Martina Fellay 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 · Nicolas Sedlatchek, Sion · Idiap, Martigny
conception graphique Formaz — Andenmatten, Sion **typographies** Cera · Franziska **papier** PlanoJet® FSC
impression MengisDruck AG, Viège **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 dynamisme au service de l'innovation et de la compétitivité
Olivier Dumas, président du Conseil de fondation de l'Idiap

05 Des étoiles européennes et valaisannes
Prof. Hervé Boulard, directeur de l'Idiap

Organisation

07 Idiap, carte d'identité

10 Thèmes de recherche, domaines d'application

11 Organigramme opérationnel

12 Groupes de recherche

14 Conseil de fondation 2015

15 Comité d'accompagnement

16 Collaborateurs

Finances

19 Compte d'exploitation

20 Sources de financement / Charges / Commentaires

21 Bilan

Recherche

23 Valais*Wallis Digital : digitaliser la mémoire collective valaisanne

25 Valais+ : une plate-forme pour la récolte de données mobiles à large échelle

26 SUMMA, MuMMER, TeSLA : succès européens

28 DexROV : développer un robot sous-marin semi-autonome

29 Youth@Night : une étude menée au travers des smartphones et des réseaux sociaux

Réseau

31 ICC'2015 : après l'incubation, l'heure de l'envol

33 MultiVEO : un système de reconnaissance vocale multilingue pour la restauration rapide

34 Transfert de technologie : faire fructifier les travaux de recherche

Visages

37 David Ginsbourger, chercheur senior :
la statistique comme langage interdisciplinaire

38 Elisa Bovio, program manager et aide-comptable :
la face cachée des projets de recherche

39 Kenneth Funes Mora, postdoctorant et entrepreneur :
suivez mon regard !

40 25^e anniversaire : 25 bougies pour l'Idiap

42 Distinctions

43 Thèses achevées

Encarté scientifique

II Idiap Research Areas : Human and Media Computing

IV Scientific Report

XXVIII Main projects in progress

XXXVI Major publications / Conferences



Un dynamisme au service de l'innovation et de la compétitivité

En route pour son 25^e anniversaire, qu'il fêtera en 2016, l'Institut de Recherche Idiap se porte à merveille. Malgré des conditions économiques devenues plus difficiles, il poursuit son développement harmonieux, avec dynamisme et efficacité.

Il appartient au Conseil de fondation de définir le cadre de travail et la mission des chercheurs afin d'assurer la pérennité de l'Idiap. Des conditions favorables — autonomie, partenariats académiques, cadre social — doivent être proposées pour prospérer de façon durable dans un rude contexte de concurrence au niveau international. Nous devons offrir la plus large liberté possible à nos scientifiques, sans perdre de vue toutefois que toute recherche doit, in fine, rester utile. Leur indépendance permet d'éviter toute récupération mercantile, où seule compterait la rentabilité immédiate, mais ils ne peuvent occulter la réalité économique. Seule la création de richesse par le « monde vrai » permet en effet de financer la recherche.

Il est admis que les entreprises qui se démarquent aujourd'hui sont celles qui misent sur la formation de leurs employés et sur l'innovation, tant au niveau de la valeur ajoutée de leurs produits que de leur conquête de marchés pour commercialiser produits et services innovants. Proche de la Fondation The Ark, chargée d'organiser et coordonner les différentes activités permettant l'établissement, l'éclosion, la croissance et l'épanouissement de sociétés en Valais, l'Idiap favorise les transferts de technologie et donne ainsi naissance à de nombreuses start-up. The Ark incarne la stratégie développée par les autorités valaisannes en vue de créer un véritable parc technologique dans le canton. Sa particularité est d'être multisite. Une de



« L'Idiap assiste de nombreuses entreprises, notamment valaisannes. »

Olivier Dumas

Président du Conseil de fondation de l'Idiap

ses antennes — IdeArk SA — est située à Martigny. Elle est animée par François Foglia, directeur-adjoint de l'Idiap, et compte actuellement 22 sociétés actives.

De surcroît, l'Idiap assiste de nombreuses entreprises, notamment valaisannes. Des partenariats sont désormais noués à dessein d'adapter les résultats de nos recherches aux besoins spécifiques de l'industrie, notamment au travers de projets financés par la Commission pour la technologie et l'innovation (CTI). De multiples projets voient le jour et permettent d'améliorer les processus de fabrication. La compétitivité de ces sociétés s'en trouve ainsi améliorée.

Les soutiens financiers de la Confédération, du Canton du Valais et de la Ville de Martigny permettent ainsi à l'Idiap de jouer un rôle actif dans l'établissement et la croissance de sociétés. Les collaborateurs de l'Idiap peuvent être fiers de participer à l'effort que notre pays doit impérativement entreprendre pour conserver une place enviable dans le concert des nations économiquement performantes. Qu'ils et elles soient ici remerciés pour leur intense labeur !

Des étoiles européennes et valaisannes

En 2014, de sombres nuages se sont accumulés dans le ciel de la recherche helvétique, laissant présager d'une année 2015 mouvementée.

Les rapports avec l'Europe s'étaient rapidement tendus après le 9 février 2014 suite à l'acceptation par le peuple de l'initiative « Contre l'immigration de masse ». Et la science a été la première à en faire les frais, à la grande surprise de tous les acteurs concernés. La cherté du franc, puis l'abandon du taux plancher face à l'euro par la BNS début 2015 plombaient durablement une économie soudainement moins encline à investir dans l'innovation. Enfin, au niveau de l'Idiap, il fallait aussi surmonter la fin du Pôle de recherche national IM2, un projet de longue durée (12 ans) qui a permis à notre institution de jouer dans la cour des grands et lui a donné une visibilité auprès du grand public.

En dépit de ces mauvais augures, 2015 s'est révélé être à nouveau un grand cru pour l'Idiap, et ce grâce à l'engagement de toutes ses collaboratrices et tous ses collaborateurs, que je profite ici de remercier. Leur travail a porté ses fruits dans tous les domaines : des projets de recherche à la création de start-up, en passant par le transfert de technologie. L'Idiap compte bien continuer dans cette voie, en exploitant pleinement son environnement unique et en gardant en tête qu'il est toujours possible de faire mieux !

« 2015 s'est révélé être à nouveau un grand cru pour l'Idiap. »

Hervé Bourlard

Professeur EPFL, directeur de l'Idiap

Les projets, justement : si l'on se réjouit légitimement du lancement de nouvelles recherches, on oublie trop souvent qu'elles ont aussi une fin ! Comme leur financement est généralement assuré pour une durée de trois à quatre ans, le taux de renouvellement est important. De plus, trouver de nouveaux collaborateurs est une tâche laborieuse. A ce titre, l'année 2015 est une réussite puisqu'elle a vu le lancement de 13 nouveaux projets. Et chose encore plus réjouissante, nos chercheurs ont réussi l'exploit de s'adjuger le financement de trois nouveaux projets européens (voir pp. 26-27) !

Au-delà de la reconnaissance internationale de l'Idiap dans ses domaines de recherche privilégiés, une des clés de cette perpétuelle chasse aux idées réside dans l'élargissement ciblé du champ de nos compétences. Celui-ci passe notamment par une collaboration accrue avec l'industrie et le recrutement de nouveaux talents. Ces deux dernières années, ce sont ainsi trois chercheurs seniors qui sont venus enrichir nos rangs.

Malgré ses succès, notre institut reste en permanence conscient de son environnement et sait qu'il doit continuer à renforcer ses liens avec ses partenaires académiques privilégiés, tant nationaux qu'internationaux. C'est pourquoi nous poursuivons aussi notre collaboration étroite avec l'EPFL, ainsi qu'avec son antenne valaisanne EPFL-Valais-Wallis au travers d'un Plan de Développement Joint. En effet, ce n'est qu'en travaillant de concert que les acteurs de la recherche et du développement permettront à la Suisse de faire face aux nombreux défis actuels et futurs. Sans la confiance et le soutien que nous accordent la Confédération, le Canton du Valais et la Ville de Martigny depuis près de vingt-cinq ans, tout ceci n'aurait pu être accompli. Qu'ils en soient ici encore remerciés !





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

Déferlement d'outils technologiques

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

Réseau national et international

Au niveau suisse, l'Idiap travaille avec les écoles polytechniques, les hautes écoles et les universités où se pratiquent également des activités de recherche. Impliqué dans plusieurs projets européens, l'institut entretient des liens étroits avec de nombreux partenaires, essentiellement en France, en Angleterre et en Allemagne. Outre-Atlantique, c'est la Californie avec l'International Computer Science Institute (ICSI) de Berkeley qui se positionne comme le partenaire privilégié.

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

47% fonds publics



53% projets de recherche

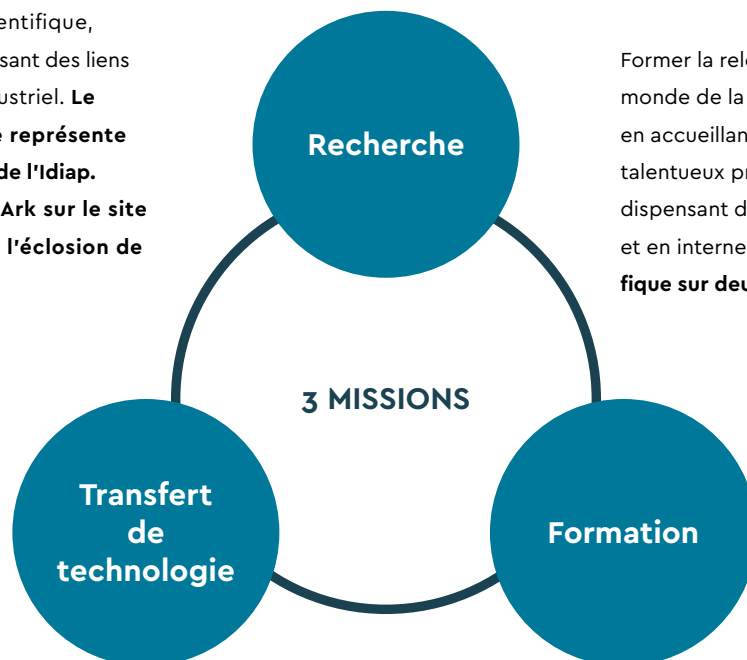
Les trois missions de l'Idiap

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 :

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.**

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-up.**

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.**



L'Idiap en chiffres (2015)

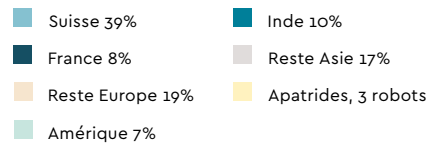
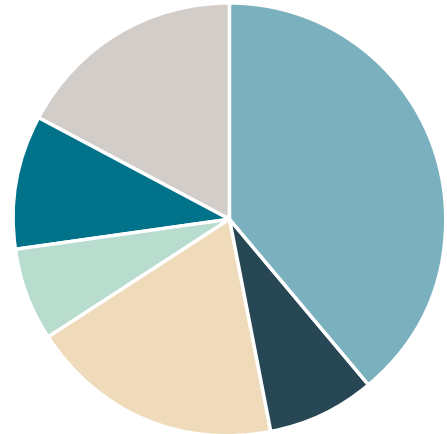
Ressources humaines

- 2 professeurs
- 2 maîtres d'enseignement et de recherche (MER)
- 9 chercheurs permanents
- 28 postdoctorants et collaborateurs scientifiques
- 31 doctorants
- 14 ingénieurs de développement
- 6 ingénieurs système
- 10 collaborateurs administratifs
- 28 stagiaires (moyenne/année)
- 2 visiteurs

132 personnes au total

99 équivalents plein temps

AU TOTAL, 28 NATIONALITÉS SONT REPRÉSENTÉES AU SEIN DE L'IDIAP

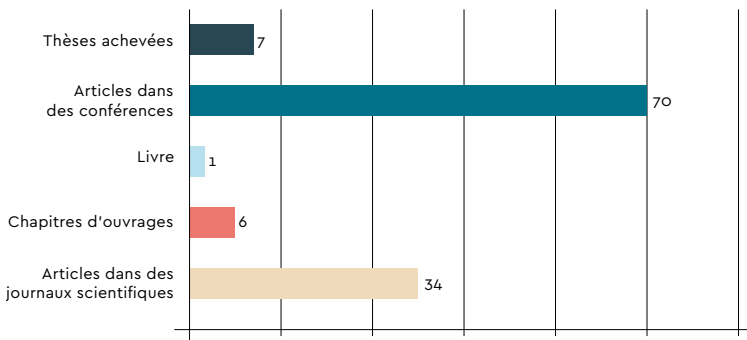


Activités scientifiques

- Participation en 2015 à 58 programmes de recherche
- Direction de projet dans plus de 25 consortiums (Suisse, Europe, projets industriels)
- Participation à la stratégie de développement économique du Canton du Valais à travers le programme The Ark et en particulier la société IdeArk
- Plus de 50 postes dans les start-up IdeArk

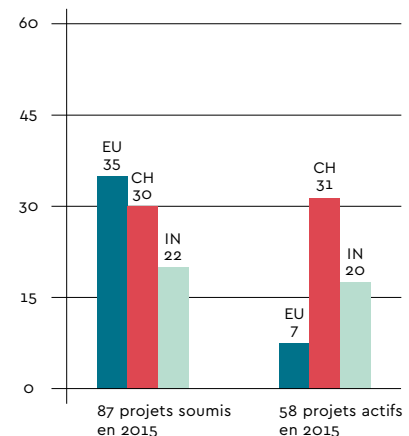
Soumission et financement des projets de recherche

En 2015, l'Idiap a soumis 87 projets (30 aux agences de financement suisses, 35 aux fonds européens et 22 projets industriels). En parallèle, l'institut a géré 58 projets actifs durant l'année.



Publications

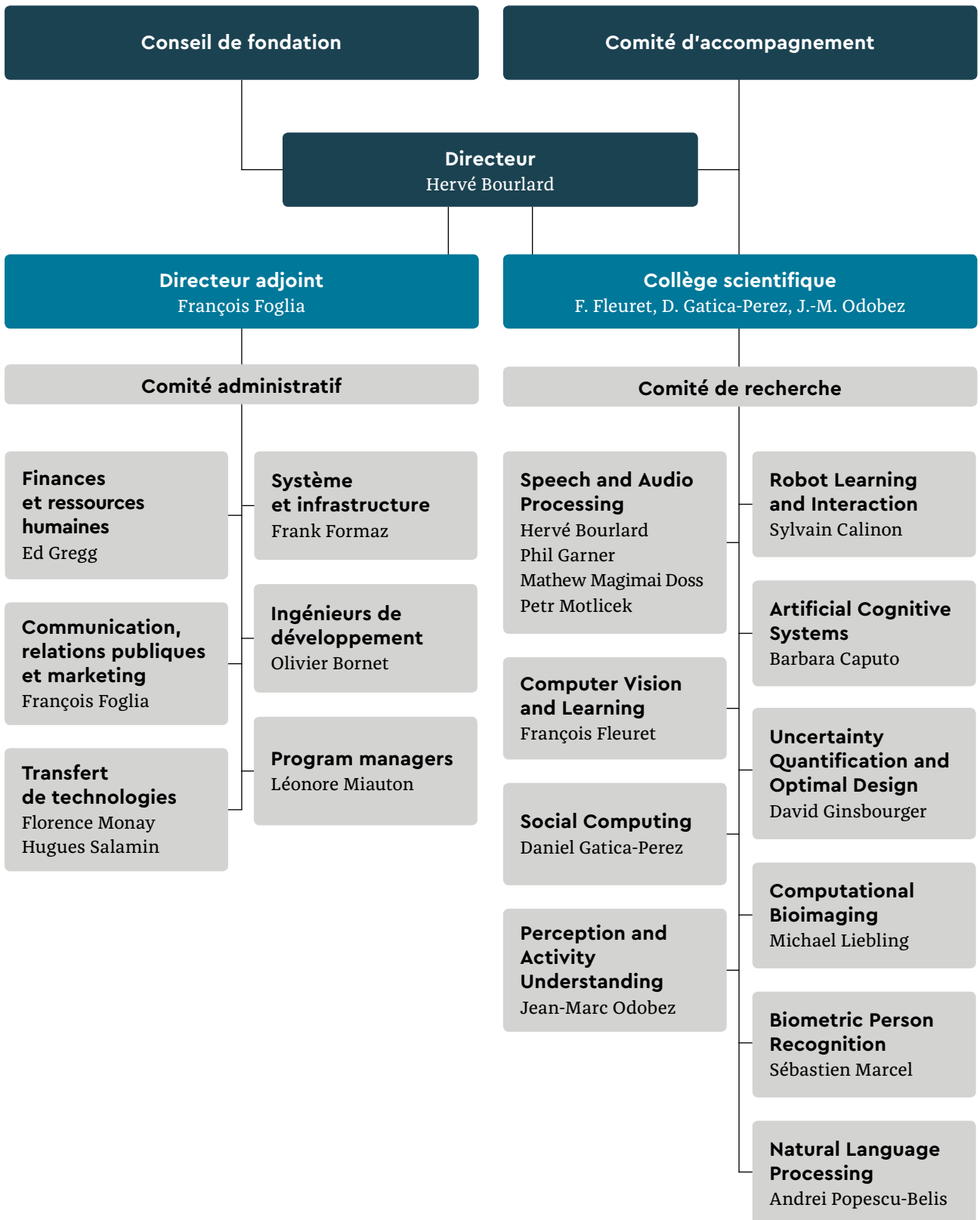
En 2015, les scientifiques de l'Idiap ont participé à 118 publications revues par les pairs.



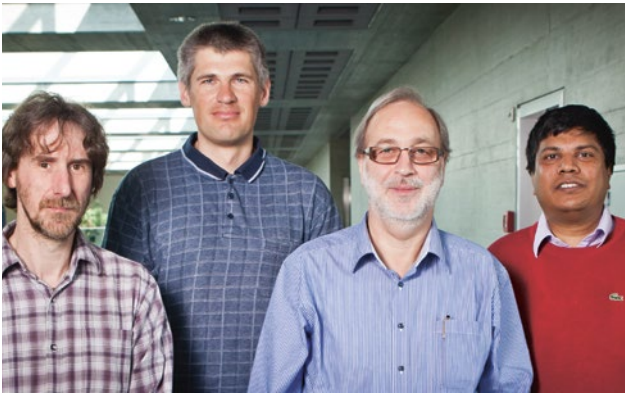
5 thèmes de recherche 10 domaines d'application



Organigramme opérationnel



Groupes de recherche



Speech & Audio Processing

Prof. Hervé Boulard (3^e depuis la gauche),
D^r Phil Garner, D^r Petr Motlicek, D^r Mathew Magimai Doss
H-index: HB: 56, PG: 20, PM: 14, MMD: 21

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).

Computer Vision & Learning

D^r 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 principales applications sur lesquelles le groupe travaille sont l'analyse automatique d'images, en particulier l'extraction automatique de sémantique, la détection d'objets, et le suivi de personnes et de structures biologiques.

À l'Idiap, les travaux sont menés au sein de dix groupes de recherche. Celui chargé du traitement de la parole et du son est emmené par le directeur de l'institut, Hervé Boulard, et trois chercheurs seniors. Les autres groupes sont également pilotés par des chercheurs seniors.

Social Computing

Prof. Daniel Gatica-Perez
H-index: 45



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.

Perception & Activity Understanding

D^r Jean-Marc Odobez
H-index: 34



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.



Robot Learning & Interaction

D^r Sylvain Calinon
H-index: 29

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.



Artificial Cognitive Systems

D^r Barbara Caputo
H-index: 32

Ce groupe travaille sur le développement d'algorithmes d'apprentissage multimodaux qui permettent aux vecteurs artificiels d'agir de manière autonome dans un cadre réaliste. Un accent particulier est mis sur la capacité de détecter de manière indépendante les lacunes de connaissance et de les combler de manière autonome avec des stratégies d'apprentissage ouvertes et flexibles. L'objectif visé est la conception d'algorithmes de principe qui sont informatiquement efficaces et fournissent une performance robuste dans un cadre très réaliste tout en donnant des garanties théoriques sur le comportement attendu.



Uncertainty Quantification and Optimal Design

D^r David Ginsbourger
H-index: 16

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



Computational Bio Imaging Group

D^r 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.



Biometric Person Recognition

D^r Sébastien Marcel
H-index: 32

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.



Natural Language Processing

D^r Andrei Popescu-Belis
H-index: 21

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



M. Jean-Daniel Antille, Vice-président
Responsable de l'Antenne Régions
Valais romand



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



Prof. Karl Aberer, Vice-Président
Vice-Président pour les systèmes
d'information, Ecole polytechnique
de Lausanne (EPFL)



M. Jordi Montserrat
Directeur régional de Venturelab



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



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



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



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



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



M. David Watrin
Responsable de l'unité de produit
«Sécurité & Intelligence» chez Swisscom

Conseil stratégique international

Le comité d'accompagnement 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. Jordan Cohen
Independent Consultant,
SPELAMODE Half Moon Bay, CA, USA



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



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



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



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



Prof. Gerhard Sagerer
Rector, Bielefeld Universität
Germany



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



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



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



Prof. Bayya Yegnanarayana
Professor and Microsoft Chair,
International Institute of Information
Technology (IIIT)
Hyderabad, India

Collaborateurs

PERSONNEL SCIENTIFIQUE

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

Oya Aran Karakus, collaboratrice scientifique, Turquie, 2009
 Afsaneh Asaei, postdoctorante, Iran, 2008
 Dario Azzimonti, doctorant, Italie, 2015
 Joan Isaac Biel, postdoctorant, Espagne, 2008
 Hervé Bourlard, directeur, Belgique/Suisse, Saxon, 1996
 Sylvain Calinon, chercheur, Suisse, Martigny, 2014
 Gulcan Can, doctorante, Turquie, 2013
 Olivier Canévet, doctorant, France, 2012
 Barbara Caputo, chercheuse senior, Italie, 2005
 Tatjana Chavdarova, doctorante, Macédoine, 2014
 Ivana Chingovska, doctorante, Macédoine, 2011
 Tiago De Freitas, doctorant, Brésil, 2014
 Subhadeep Dey, doctorant, Inde, 2014
 Pranay Dighe, doctorant, Inde, 2013
 Trinh-Minh-Tri Do, collaborateur scientifique, Vietnam, 2009
 Elie El Khoury, postdoctorant, Liban, 2011
 Laurent El Shafey, doctorant, France, 2010
 Marc Ferras Font, collaborateur scientifique, Espagne, 2011
 François Fleuret, chercheur senior, France, 2007
 Kenneth Funes Mora, postdoctorant, Costa Rica, 2011
 Philip Garner, chercheur senior, Angleterre, 2007
 Daniel Gatica-Perez, chercheur senior, Mexique, 2002
 Sucheta Ghosh, postdoctorante, Inde, 2014
 David Ginsbourger, chercheur senior, France, 2015
 Manuel Günther, postdoctorant, Allemagne, 2012
 Ioannis Havoutis, postdoctorant, Grèce, 2015
 Alexandre Heili, postdoctorant, France, 2010
 Guillaume Heusch, collaborateur scientifique, Suisse, Lausanne, 2015
 Ivan Himawan, postdoctorant, Australie, 2014
 Pierre-Edouard Honnet, doctorant, France, 2012
 Rui Hu, postdoctorante, Chine, 2013
 David Imseng, postdoctorant, Suisse, St. German, 2009
 Cijo Jose, doctorant, Inde, 2013
 Pavel Korshunov, postdoctorant, Estonie, 2015
 Tipaluck Krityakierne, postdoctorante, Thaïlande, 2015
 Ilja Kuzborskij, doctorante, Lituanie, 2012
 Alexandros Lazaridis, collaborateur scientifique, Grèce, 2012
 Nam Le, doctorant, Vietnam, 2015
 Rémi Lebret, doctorant, France, 2012
 Joël Legrand, doctorant, France, 2012

Michael Liebling, chercheur senior, Suisse, St-Sulpice, 2014
 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, France/Suisse, Martigny, 2000
 Petr Motlicek, chercheur, République tchèque, 2005
 Hannah Muckenhirn, doctorante, France, 2015
 Skanda Muralidhar, doctorant, Inde, 2014
 James Newling, doctorant, Angleterre, 2013
 Laurent Nguyen, postdoctorant, Suisse, Lausanne, 2011
 Jean-Marc Odobez, chercheur senior, France/Suisse, Clarens, 2001
 Pedro Oliveira Pinheiro, doctorant, Brésil, 2012
 Dimitri Palaz, doctorant, Suisse, Martigny, 2011
 Nikolaos Pappas, doctorant, Grèce, 2012
 Novi Patricia, doctorante, Indonésie, 2012
 Trung Phan, doctorant, Vietnam, 2015
 Andrei Popescu-Belis, chercheur senior, France/Roumanie/Suisse, Lausanne, 2007
 Blaise Potard, postdoctorant, France, 2013
 Xiao Pu, doctorante, Chine, 2014
 André Rabello Dos Anjos, collaborateur scientifique, Brésil, 2010
 Dhananjay Ram, doctorant, Inde, 2014
 Ramya Rasipuram, postdoctorante, Inde, 2010
 Marzieh Razawi, doctorante, Iran, 2013
 Lakshmi Saheer, collaboratrice scientifique, Inde, 2008
 Darshan Santini, doctorant, Inde, 2012
 Ajay Tanwani, doctorant, Pakistan, 2015
 Pedro Tome, postdoctorant, Espagne, 2014
 Raphaël Ullman, doctorant, Suisse, Lausanne, 2012
 Matthias Vanoni, doctorant, France, 2013
 Di Wu, postdoctorant, Chine, 2015
 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

Frédéric Dubouchet, ing. de développement, Suisse, Lens, 2014
Charles Dubout, ing. de développement, France, 2015
Samuel Gaist, ing. de développement, Suisse, Ardon, 2013
Salim Kayal, ing. de développement, Suisse, Vevey, 2011
Vasil Khalidov, ing. de développement senior, Russie, 2010
Andreas Kobler, ing. de développement, Suisse, 2015
Christine Marcel, ing. de développement, France/Suisse,
Martigny, 2007
Florent Monay, officier de transfert de technologie, Suisse, Choëx, 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
Sergio Calabretta, relations industrielles, Italie, 2014
Antoine Dorsaz, aide-comptable, Suisse, Fully, 2012
Christophe Ecoeur, program manager, Suisse, Collombey, 2010
Martina Fellay, program manager, Autriche, 2012
François Foglia, adjoint du directeur, Suisse, Saxon, 2006
Edward-Lee Gregg, responsable financier, Etats-Unis, 2004
Léonore Miauton, responsable program managers, Suisse,
Chexbres, 2012
Sylvie Millius, assistante administrative, Suisse, Vétroz, 1996
Nadine Rousseau, assistante administrative, Belgique/Suisse,
Saxon, 1998

• Ingénieurs système

Bastien Crettol, ingénieur système, Suisse, Sion, 2005
Norbert Crettol, ingénieur système, Suisse, Martigny, 2002
Cédric Dufour, ingénieur système, Suisse, Aigle, 2007
Frank Formaz, administrateur système, Suisse, Fully, 1998
Louis-Marie Plumel, ingénieur système, 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.

Christian Abbet, Suisse, EPFL, Lausanne
Afroze Baqapuri, Pakistan, EPFL, Lausanne
Daniel Berio, Italie, University of London (GBR)
Sushil Bhattacharjee, Suisse, Talaris Ltd (GBR)
Fabian Brix, Allemagne, EPFL, Lausanne
Kevin Chan, Etats-Unis, University of California, Santa Barbara (USA)
Gilberto Chávez-Martínez, Mexique, CIMAT, Guanajuato (MEX)
Yiqiang Chen, Chine, INSA Lyon (FRA)
Artur Costa Pazo, Espagne, Universidade de Vido (ESP)
Ailbhe Finnerty, Irlande, University of Trento (ITA)
Branislav Gerazov, Macédoine, University of Skopje (MKD)
Nicolas Gninenko, Suisse, EPFL, Lausanne
David Guennec, France, Université Rennes 1 (FRA)
Maryam Habibi, Iran, doctorante Idiap
Serife Kucur Ergunay, Turquie, EPFL, Lausanne
Jeevanthi Uthpala Liyanapathirana, Sri Lanka, Copenhagen Business School (DNK)
Gil Luyet, Suisse, Université de Fribourg
Lesly Miculicich, Pérou, Université de Fribourg
Terry Niederhauser, Suisse, HESSO
Emmanuel Pignat, Suisse, EPFL, Lausanne
Dairazalia Sanchez-Cortes, Mexique, University Ensenada of Baja California (MEX)
Ashtosh Sapru, Inde, doctorant Idiap
Milan Secujski, Serbie, Université de Novi Sad (SRB)
Alexandre Sierro, Suisse, HESSO
Dominique Thao, Suisse, EPFL, Lausanne
Julien Theux, Suisse, collège de Brigue
Bogdan Vlasenko, Ukraine, Otto-von-Guericke Universität, Magdebourg (GER)
Sree Harsha Yella, Inde, doctorant Idiap

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.

Nilceu Marana, Brésil, Université de Sao Paulo (BRA)
Shogo Okada, Japon, Tokyo Institute of Technology (JPN)



finances

18 — 21

Compte d'exploitation

PRODUITS	2014	2015	%
Commune de Martigny	700 000	700 000	6.8%
État du Valais	1 720 000	1 720 000	16.7%
Confédération	2 550 000	2 467 300	23.9%
Loterie Romande	50 000	100 000	1%
Projet PRN IM2	476 648	-	-
Projets Fonds national suisse de la recherche scientifique	1 427 889	1 715 249	16.5%
Projets Fondation Hasler	371 045	606 689	5.9%
Projets européens	1 210 890	1 395 595	13.5%
Projets The Ark	173 797	225 070	2.2%
Projets CTI	607 217	388 170	3.8%
Contribution EPFL	72 000	72 000	0.7%
Financement industriel	366 736	538 818	5.2%
Autres financements / Produits exceptionnels	685 926	396 274	3.8%

TOTAL DES PRODUITS	10 412 148	10 325 165	100.00 %
---------------------------	-------------------	-------------------	-----------------

CHARGES

Frais de personnel	7 879 097	7 884 132	76.4%
Formation et déplacements	355 766	371 774	3.6%
Partenaires externes	302 540	167 114	1.6%
Informatique : matériel et maintenance	246 284	185 680	1.8%
Frais administratifs	233 001	148 148	1.5%
Promotion et communication	50 079	45 468	0.5%
Loyer et charges	818 888	791 284	7.7%
Amortissement	310 853	263 690	2.5%
Charges exceptionnelles	-	-	-
Provisions	200 000	455 000	4.4%

TOTAL FONDS ÉTRANGERS	10 396 508	10 312 290	99.88 %
------------------------------	-------------------	-------------------	----------------

RÉSULTAT D'EXPLOITATION	15 640	12 875	0.12 %
--------------------------------	---------------	---------------	---------------

Commentaires sur les comptes 2015

Idiap réalise un excellent exercice 2015. Et ceci, malgré l'effet dû à l'abandon du taux plancher face à l'euro. Les comptes annuels dépassent, une nouvelle fois, la barre des 10 millions de francs.

La part des subventions publiques respecte toujours l'objectif stratégique fixé par la Direction de se situer en-dessous des 50%. Pour 2015, le ratio est d'environ 47%.

Depuis quelques années, l'Idiap met l'accent sur le développement de ses relations industrielles cantonales, nationales et internationales. Force est de constater que ces efforts portent leurs fruits. Plus de 10% du budget est, aujourd'hui, consacrés à des projets en lien avec l'économie (financement TheArk, CTI, industriel).

Les frais de personnel (salaire + charges sociales) forment la majeure partie des charges, soit un peu moins de 80%.

L'Idiap boucle ses comptes 2015 sur un excédent de 12'875 francs.

Subventions

Confédération, Canton, Commune

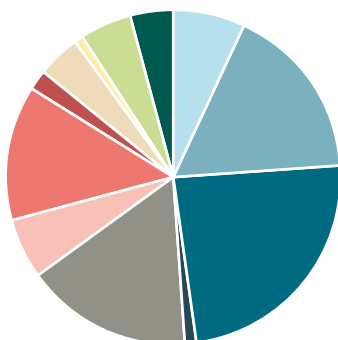
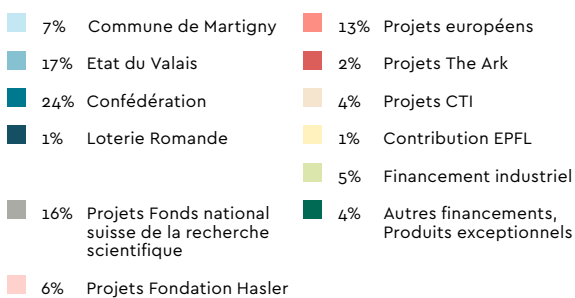
(En milliers de francs suisses)

ANNÉES	2013	2014	2015	2016*
Confédération	2400	2550	2467	2424
Canton	1320	1720	1720	1720
Commune	700	700	700	700

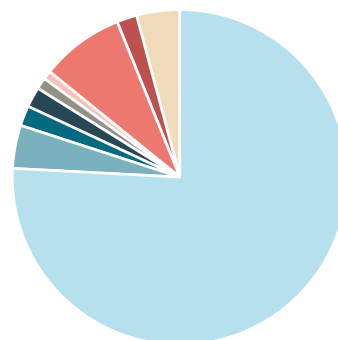
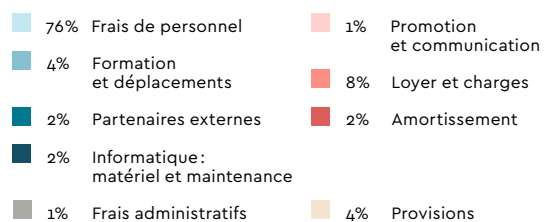
20

*Budget

Répartition des sources de financement



Répartition des charges




Bilan (CHF)

ACTIFS	31.12.2014	31.12.2015
Liquidités	3 172 241.55	3 467 573.13
Débiteurs	315 159.11	319 121.25
Actifs de régularisation et divers	502 413.88	622 013.91
TOTAL ACTIFS CIRCULANTS	3 989 814.54	4 408 708.29
Mobilier et matériel informatique	515 818.64	434 031.64
Participations financières	10 000.00	10 000.00
TOTAL ACTIFS IMMOBILISÉS	525 818.64	444 031.64
TOTAL ACTIFS	4 515 633.18	4 852 739.93
PASSIFS	31.12.2014	31.12.2015
Créanciers	300 607.00	349 413.85
Passifs de régularisation	2 058 214.07	1 878 638.93
Provisions	950 000.00	1 305 000.00
TOTAL FONDS ÉTRANGERS	3 308 821.27	3 533 052.78
Capital	40 000.00	40 000.00
Réserve spéciale	1 000 000.00	1 100 000.00
Résultat reporté	151 171.95	166 812.11
Résultat net	15 640.16	12 875.04
TOTAL FONDS PROPRES	1 206 812.11	1 319 687.15
TOTAL PASSIFS	4 515 633.18	4 852 739.93



$p(s|M)$ $p(x|M)$

$h(n):$ 

$$f(n) = g_0 + \sum_{i=1}^K \left(\int_{\tau} g_i(\tau) x_i(n-\tau) d\tau \right)$$

$$f(n) = \sum_{i=-50}^{50} h(i) y(n-i), \quad n: 0 \text{ to } K$$

$$\sum_{n=0}^K f(n) y(n-i)$$

$$\sum_{i=-50}^{50} h(i) y(n-i)$$
$$\sum_{j=n-50}^{50} h(j) x(i) + \dots$$

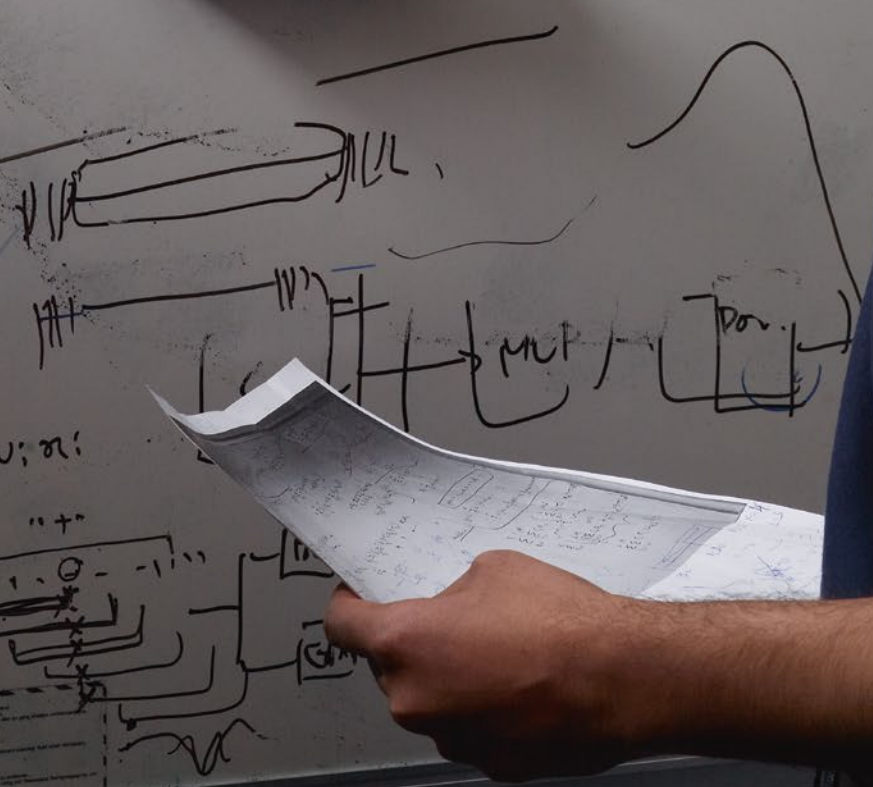
$$x_0 h_{-50} + x_0 h_{-49} \dots x_0 h_0 + x_1 h_1 \dots$$

$$x_1 h_{-50} + \dots + x_0 h_1 + x_1 h_1$$

$$\sum_{i=-50}^0 h_i + \sum_{i=1}^K h_i$$

recherche

22 — 29



Digitaliser la mémoire collective valaisanne

En 2015, le Canton du Valais a fêté le 200^e anniversaire de son entrée dans la Confédération helvétique. Afin de commémorer ce bicentenaire, l'Etat du Valais a lancé un appel à idées auquel l'Idiap et ses partenaires ont répondu en soumettant le projet Valais*Wallis Digital. Son but était de profiter des possibilités offertes par les nouvelles technologies pour digitaliser la mémoire collective du Vieux-Pays.

La mémoire collective se compose bien évidemment de documents d'archives des institutions publiques, mais aussi des souvenirs d'associations et de personnes privées. Isolés, ces derniers ont avant tout une valeur personnelle. Ce n'est qu'une fois rassemblés, contextualisés et mis en lien qu'ils prennent toute leur valeur. Or bien souvent, ces archives privées dorment dans une cave ou un grenier. Pourtant, elles sont une mine d'or pour illustrer une histoire sociale et culturelle d'une ville, d'une région ou d'un canton.

Collecter et partager les souvenirs

À l'occasion du bicentenaire de l'entrée du Valais dans la Confédération, l'Idiap a donc lancé avec ses partenaires (la Médiathèque Valais, les Archives de l'Etat du Valais, la Coopérative Migros Valais et l'Ecole professionnelle des arts contemporains) le projet Valais*Wallis Digital dans le but d'offrir à la population valaisanne une plate-forme lui permettant de déposer et rassembler sous forme digitale les photos, films, fichiers audio ou documents écrits à valeur historique qu'elle possède à domicile. La plate-forme ne sert bien évidemment pas qu'à récolter les archives mais aussi à les partager. Elle est ainsi toujours accessible à toutes et à tous à l'adresse www.valais-wallis-digital.ch.



« Un tel projet sort clairement de ce que nous avons l'habitude de faire. »

François Foglia

Directeur adjoint, Idiap

Mais participer à la sauvegarde du patrimoine valaisan permet également de contribuer aux recherches en sciences digitales de l'Idiap. Ces dernières consistent notamment à traiter et indexer de larges bases de données multimédias résultant d'une démarche participative de la population. Ce type de projets a un impact qui va bien au-delà du projet Valais*Wallis Digital et vise à développer de nouveaux outils informatiques permettant de faire face aux grands défis sociétaux à venir, tels que la gestion de l'énergie, des soins de santé et de l'environnement.

Jeu de cartes

Parallèlement, les partenaires du projet ont créé, dans le cadre des Mania de la Migros, un jeu de cartes illustrant 200 événements marquants du Valais des deux derniers siècles, sur des thèmes aussi variés qu'agriculture et gastronomie, art, culture et spiritualité, générations, histoire et tradition, immigration et émigration, innovation et technologie, musique et médias, sport, tourisme, etc.



24

Réaliser ce jeu de cartes a été un travail de longue haleine. Pour François Foglia, directeur adjoint de l'Idiap et responsable du projet Valais*Wallis Digital, « un tel projet sort clairement de ce que nous avons l'habitude de faire, que ce soit le choix graphique, de la taille des cartes, des effets visuels associés aux jokers, ou de la logistique de distribution des cartes. J'en ai pleinement pris conscience lorsque j'ai vu les 33 palettes remplies de cartes dans les entrepôts de la Migros ! ».

À l'heure du bilan, François Foglia affiche une satisfaction en demi-teinte : « Pour ce qui est de la plate-forme, la technique fonctionne bien. Mais le nombre de documents déposés est en dessous de nos attentes. Nous aurions peut-être dû nous focaliser sur des thèmes précis. D'ailleurs, notre récent appel sur les archives en lien avec les guggens et les fanfares a rencontré plus de succès : le cercle des personnes intéressées est plus restreint, et ses membres plus enclins à fouiller leurs archives pour les partager avec d'autres passionnés. Mais, dans une logique scientifique, ce constat permet aussi de lancer de nouvelles recherches, comme le projet Valais+ (voir l'article p.25), dont un volet aborde la question de comment inciter un public à participer à une collecte de données numériques thématiques. »

Valais+ : Une plate-forme pour la récolte de données mobiles à large échelle

Au travers du projet Valais+, financé par la Loterie Romande, l'Idiap entend mettre à disposition de la population suisse une plate-forme internet pour la récolte de données mobiles à large échelle. Basée sur le concept de défis, cette plate-forme repose sur l'initiative de personnes individuelles ou de collectivités qui désirent récolter des données. Elle permettra de créer un défi, de le personnaliser, et de le rendre disponible sur l'application mobile Civique.

Le smartphone permet de récolter très facilement des images, des vidéos, des textes et de les géolocaliser. Si une récolte individuelle est possible avec les applications de base, il en va tout autrement du partage et de la structuration de ces informations.

Innovation sociale

À l'aide de subsides de la Loterie Romande, le professeur Daniel Gatica-Perez (responsable du groupe « Social Computing ») a décidé de se lancer dans un projet d'innovation sociale avec en arrière-pensée de mettre le savoir-faire de l'Idiap au service des citoyens suisses et valaisans. Pour ce faire, il dirige avec le D^r Joan-Isaac Biel le développement d'une plate-forme internet versatile à l'intention des collectivités publiques, des clubs et des associations afin de récolter aisément des informations les intéressant.

« Imaginons qu'une commune souhaite régler un problème d'urbanisme, comme le réaménagement d'un parc, explique Daniel Gatica-Perez. Celle-ci pourra créer un défi sur la plate-forme et inviter ses habitants à photographier des éléments du parc actuel qu'ils jugent

insatisfaisants et de répondre à des questions prédéfinies — une place de jeu qu'ils estiment vétuste par exemple. Une forme de démarche participative que les autorités pourront mettre à profit. Ces données géolocalisées seront disponibles sur une page web dédiée, sous différentes formes de visualisation. ».

Comment motiver les utilisateurs?

Les créateurs d'un défi en seront les promoteurs. À eux de partager les informations et les motivations à l'origine de leur défi pour qu'un maximum de personnes concernées y participe. Une page web unique pour chaque défi et un mécanisme de synchronisation des défis sur l'application mobile permettront une participation et un partage simples.

Le patrimoine culturel pour pilote

Daniel Gatica-Perez souhaite débiter par un petit nombre de projets pilotes. La professeure Maria Sokhn de la HES/SO Sierre a par exemple été contactée pour la création d'un défi dans le cadre de son projet City-Zen. Pour la chercheuse, l'information touristique sur le patrimoine culturel valaisan est très large, mais aussi très diffuse. A travers son projet, elle souhaite offrir une plate-forme de recherche d'information spatio-temporelle valorisant ce patrimoine et l'information y relative. La mise sur pieds d'un défi lié au patrimoine culturel valaisan permettra de mettre à contribution les citoyens valaisans pour élargir cette base de connaissances.

Succès européens

A lors que le taux de succès de la procédure de soumission du programme-cadre pour la recherche et l'innovation de la Communauté européenne Horizon 2020 (H2020) tourne autour de quatorze pour cent environ, les chercheurs de l'Idiap ont réussi en 2015 le tour de force de voir être financés trois des onze projets qu'ils ont co-proposés dans le cadre du volet dédié aux technologies de l'information et de la communication (TIC). De la robotique de divertissement à l'analyse automatisée de données multimédia en passant par la biométrie, ces trois projets correspondent à un montant de 22 mio EUR, dont un peu plus de 3,8 mio CHF reviennent directement à l'Idiap. Ce succès éclatant contraste avec une année européenne 2014 en demi-teinte.



Analyser les flux multimédia : Scalable Understanding of Multimodal Media (SUMMA)

Un énorme flux d'information arrive en tout temps dans les bureaux des rédactions et sert de matière première pour la production de nouveaux contenus reprenant, agrégeant, complétant ou commentant ce qui existe déjà afin d'en donner le contexte. Ainsi, la diffusion des images d'un événement peut donner naissance à des articles publiés sur des plates-formes d'information, qui généreront des tweets qui pourront être repris sur la plates-forme et donner un nouvel angle d'attaque au sujet qui sera tourné pour le téléjournal du soir, peut-être agrémenté d'images d'archives. Afin de suivre ce fil d'information, des employés des grands médias passent leurs journées à indexer manuellement les contenus multimédias qu'ils analysent. Une tâche encore plus complexe si l'on songe qu'il faut l'effectuer dans plusieurs langues.

Le projet SUMMA a précisément pour but d'automatiser cette tâche afin de pouvoir encore plus vite et plus efficacement voir émerger des nouvelles ou des personnalités, faire des liens entre elles et identifier ce qui est véritablement nouveau. Dans ce projet, les chercheurs

de l'Idiap, emmené par le Prof. Hervé Bourlard, travaillent sur la transcription et l'indexation du contenu audio en neuf langues. « A l'Idiap, nous savons bien le faire pour l'anglais, le français et l'allemand. Mais là, nous aurons aussi affaire au mandarin, au perse ou encore au letton », explique-t-il. Ils devront également lier ces contenus aux fils d'actualités de réseaux sociaux comme Facebook ou Twitter. Pour Hervé Bourlard, un autre défi est la quantité d'information à traiter : « Les partenaires médias du projet — BBC, Deutsche Welle, Latvian News Agency et TV Qatar — nous demandent de pouvoir traiter quotidiennement 1.5 TB, soit l'équivalent de 500 heures audio-vidéo de contenu par jour ! ». Coordonné par l'Université d'Edimbourg, le projet européen H2020 SUMMA occupera 4-5 personnes à plein temps durant trois ans à l'Idiap.

www.summa-project.eu/



Un robot pour divertir : MultiModal Mall Entertainment Robot (MUMMER)

Développer des robots capables d'interagir de manière sociale avec les clients d'une grande surface commerciale : tel est l'objectif principal du projet MUMMER, mené par un consortium de sept institutions et coordonné par l'Université de Glasgow. Les interactions planifiées visent tout d'abord à divertir le public — raconter des histoires, coordonner des jeux sociaux comme des quizz — mais sont également amenées à prendre des formes plus commerciales comme effectuer des sondages ou proposer des offres spécifiques aux consommateurs en fonction de leurs intérêts. Le robot ne se contentera donc pas d'attendre que l'on s'adresse à lui, mais pourra également adopter un comportement plus proactif en approchant les clients. Cela requerra donc d'avoir une perception sociale assez fine des usagers potentiels afin de pouvoir détecter rapidement si une interaction est désirée ou non. Histoire de compliquer encore un peu plus les choses, ces interactions pourront se faire avec des groupes et se dérouleront dans un environnement non confiné (les tests auront lieu dans un shopping mall finlandais comptant pas moins de 3000 enseignes commerciales).

Du côté de l'Idiap, le projet est aux mains du Dr Jean-Marc Odobez (en collaboration avec le Dr Petr Motlicek pour la partie audio): « Dans le consortium, nous sommes responsables de la partie perception des personnes, comprenant notamment la localisation, la prise de parole, et l'extraction des indices non-verbaux (signaux sociaux). Par rapport à notre savoir-faire, MuMMER gagne en complexité. Tant le robot que l'utilisateur pourront être perturbés par ce qui se passe aux alentours. Lorsque le robot se dirigera vers quelqu'un, d'autres personnes pourront passer entre les deux, le faisant dévier de son but si le système d'identification n'est pas suffisamment robuste. De même, le robot pourrait prendre pour un non les rotations de la tête d'un usager dont l'attention serait captée par une annonce audio. Avec les groupes, il devra maintenir une représentation continue de ses membres et de leurs activités malgré son champ de vision limité, les gestes et activités qu'il doit produire afin d'avoir un comportement social et qui affectent sa vision ou son audition, et la complexité des interactions. De manière générale, son lexique de perception devra être beaucoup plus robuste, plus complet, et plus fin ».

Enfin, ce projet de recherche sera l'occasion pour Jean-Marc Odobez de retrouver un ancien collègue de l'Idiap, en la personne d'Alessandro Vinciarelli, aujourd'hui chercheur à l'Université de Glasgow et responsable dans MuMMer de la partie synthèse de signaux sociaux perçus positivement par le public.

www.mummer-project.eu/



La biométrie pour s'identifier lors du passage d'un examen : Adaptive Trust-based e-assessment online at the different learning stages (TeSLA)

Les TIC offrent une opportunité pour s'affranchir des limitations physiques et financières qu'impose l'enseignement traditionnel. La révision de ce modèle, qui voit étudiants et corps professoral se côtoyer, exige de nouveaux standards d'identification, notamment pour tout ce qui a trait au passage d'examens. Comment garantir de façon simple et fiable que c'est bien la personne inscrite qui passe l'épreuve attestant de l'acquisition d'un savoir donné?

Voici précisément la question centrale du projet européen TeSLA, coordonné par la Fundació per a la Universitat Oberta de Catalunya. C'est grâce à l'expertise en matière de biométrie que les chercheurs de l'Idiap ont acquise, notamment lors de projets FP7 et du FNS et de collaborations industrielles, avec Google par exemple, que les membres du consortium TeSLA ont invité le Dr Sébastien Marcel à se joindre à eux. Ce projet occupera deux chercheurs à plein temps sur deux ans, ainsi qu'un ingénieur durant six mois. Ils s'attèleront à développer les méthodes de biométrie vocale et à co-développer celles de biométrie faciale. En outre, les chercheurs de l'Idiap sont en charge de la détection des attaques.

Pour Sébastien Marcel, responsable du projet à l'Idiap, « ce projet est une chance d'appliquer des techniques que nous avons développées ces dernières années. Mais les problèmes qui ne manqueront pas de se poser nous permettront à n'en pas douter de poursuivre nos recherches sous un angle plus fondamental. »

www.tesla-project.eu/

Développer un robot sous-marin semi-autonome

Le fond des océans est un endroit aussi mystérieux qu'inhospitalier. Les conditions qui y règnent rendent difficiles les pourtant très nombreuses opérations industrielles ou scientifiques (biologie, géologie, archéologie) qui s'y déroulent. Pour les effectuer, les plongeurs s'avèrent bien souvent indispensables. Mais face à la dangerosité et à la complexité des missions qu'ils effectuent, le recours à des robots (semi-)autonomes ouvre des pistes intéressantes. C'est précisément au développement d'un tel robot que travaillent les chercheurs du projet DexROV.

Les fonds marins représentent environ deux tiers de la superficie de notre planète. Pourtant, une grande partie d'entre eux nous reste plus inconnue que la surface d'étoiles lointaines. La faute à un environnement dangereux, complexe et agressif du fait de l'obscurité, de la pression, de la température et de la salinité qui y règnent.

Paradoxalement, ces lieux inhospitaliers regorgent d'infrastructures de télécommunications, d'exploitation de matières premières et de production d'énergie. Leur installation, leur maintenance et leur démantèlement sont des tâches très lourdes. Ces travaux voient bien souvent l'intervention de robots. Mais ceux-ci se limitent en général à des missions d'observation et, au final, ce sont bien souvent des plongeurs qui doivent effectuer manuellement les tâches les plus complexes. Or emmener des plongeurs à grande profondeur est chose compliquée car requérant de très longs paliers de décompression, d'une durée allant jusqu'à dix jours. De plus, devant la dangerosité du métier, plus à comparer à celle d'un mercenaire qu'à celle d'un ouvrier, il n'est pas toujours facile de trouver des plongeurs.

Les limites des opérations à distance

Quant aux robots télécommandés, ils nécessitent la présence in situ de toute une équipe de spécialistes dédiés à une seule mission bien précise. Une alternative serait d'opérer les robots à distance. Mais cette option se heurte à une limitation d'ordre physique : la transmission par satellite des informations nécessaires aux mouvements du robot impose un délai d'environ une seconde entre l'envoi d'un ordre et la réception du signal de retour, ce qui nécessite de séquencer les mouvements complexes, rallongeant d'autant leur durée.

A cet égard, le groupe Robot Learning & Interaction du Dr Sylvain Calinon travaille, dans le cadre du projet européen DexROV, sur le développement d'un robot semi-autonome qui serait précisément capable d'opérer de manière autonome durant cette seconde de latence nécessaire à la transmission des ordres. Ayant appris à effectuer divers types de mouvements, il pourrait les reconnaître et en quelque sorte les terminer ou les prolonger durant les interruptions de signal. Dans le cas

d'une vanne à ouvrir, par exemple, une fois le mouvement de rotation du volant initié, le robot devrait être à même de le poursuivre jusqu'à l'ouverture complète.

Il s'agit d'apprendre des séquences de base au robot qu'il adaptera et assemblera lui-même.

Gestuelle et briques Lego

Une période d'apprentissage est cependant nécessaire, et celui-ci peut se pratiquer de deux manières. Le robot peut assimiler l'information via une caméra qui lui montre le mouvement à effectuer, ou bien on peut littéralement lui prendre la main et le guider, un peu comme un ergothérapeute, dans la réalisation d'un geste. « Au contraire d'un robot industriel programmé pour effectuer de manière répétitive des tâches parfaitement identiques, explique Sylvain Calinon, responsable du projet à l'Idiap, il s'agit là d'apprendre des séquences de base que le robot adaptera et assemblera lui-même pour arriver à une séquence complète qui ne sera jamais deux fois la même. »

Les chercheurs de l'Idiap ne disposant pas de robot sous-marin, ils travaillent sur un robot industriel, nommé Baxter, possédant deux bras articulés contrôlés avec des commandes en force. En 2018, à l'issue du projet débuté en mars 2015 et coordonné par la société belge Space Applications Services, les participants au projet DexROV, espèrent avoir développé un démonstrateur opérationnel qu'ils auront testé à une profondeur de 1300 mètres.

Youth@Night: une étude menée au travers des smartphones et des réseaux sociaux

Employer les technologies mobiles et les réseaux sociaux et leur mode de fonctionnement pour étudier les habitudes des jeunes en matière de consommation d'alcool et de vie nocturne. Tel est le but du projet de recherche Youth@Night auquel participe le professeur Daniel Gatica-Perez.

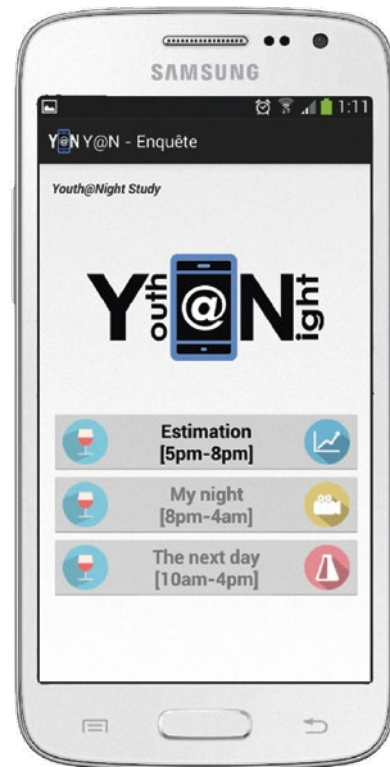
Pour être efficace, une stratégie de prévention se doit de connaître très précisément sa cible. Mais les données sur les habitudes de consommation d'alcool des jeunes adultes en Suisse sont limitées, car difficiles à obtenir. Elles proviennent généralement de questionnaires remplis a posteriori.

Mieux comprendre la dynamique nocturne

Lancé par Addiction Suisse, le groupe de Social Computing de l'Idiap et le Département de géographie de l'Université de Zurich, le projet Youth@Night cherche à mieux comprendre l'utilisation de l'espace public, à analyser l'influence du contexte au cours de la soirée et étudier le rôle de la structure urbaine sur les habitudes de consommation tant au niveau individuel qu'au niveau collectif.

Une technologie en phase avec les pratiques actuelles

Afin de mieux comprendre la problématique et de récolter des données de manière non intrusive pour les participants à l'étude, l'Idiap a développé en 2014 une application pour téléphone portable. « Notre idée est d'utiliser une technologie en phase avec les pratiques des jeunes de 16 à 25 ans, explique Daniel Gatica-Perez, responsable du projet à l'Idiap. Avec leur smartphone



et les réseaux sociaux, ils ont l'habitude de poster des informations sur leurs activités et l'endroit où elles se déroulent. » Tout au long de la soirée, l'application Y@N demande des informations sur les quantités d'alcool ingérées et les lieux fréquentés. L'utilisateur doit également prendre des photos de ses boissons et de courtes vidéos indiquant le contexte de consommation. « Ces petites fenêtres sur la vie nocturne servent également de base de discussion pour une série d'entretiens qualitatifs qui seront menés par la suite », poursuit le chercheur.

Les données fournies par les participants à l'étude — une centaine à Zurich et une autre à Lausanne — sont agrégées puis comparées à des données de l'application « Foursquare », avec laquelle les gens renseignent d'autres utilisateurs sur leurs positions, afin de voir si des tendances communes y émergent. A noter que les deux jeux de données sont anonymisés.

« A l'aide de « Foursquare », nous avons également établi une carte des hot spots de la vie nocturne lausannoise et zurichoise. Elle nous a permis de mieux localiser notre public cible pour y recruter des volontaires », conclut Daniel Gatica-Perez.

réseau

30 — 35

Après l'incubation, l'heure de l'envol

L'International Create Challenge (ICC), lancé par l'Idiap en 2012, 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 2015, ce programme unique a une nouvelle fois connu un beau succès. Trois projets susceptibles, demain, d'améliorer notre quotidien ont ainsi été distingués.

Du 25 août au 15 septembre 2015, l'Idiap a accueilli sept équipes d'entrepreneurs en herbe — scientifiques, designers ou entrepreneurs — en provenance de Suisse, de France, d'Espagne, de Roumanie et même du Costa Rica. A l'issue de l'ICC'2015, le jury, composé de personnalités du monde de la recherche, du capital-risque et de l'innovation, a distingué les trois projets alliant innovation et haute qualité scientifique que sont Eyeware (voir article p. 39), Dialexy et App4autism.

EYEWARE*

ÉQUIPE

- Dr Kenneth Funes Mora : postdoctorant à l'Idiap
- Carlos Becker : doctorant dans le domaine Computer Vision à l'EPFL
- Serban Mogos : doctorant dans le domaine Technological Change & Entrepreneurship à Carnegie Mellon University (USA)

OBJECTIF

Développement de solutions qui donnent la capacité aux ordinateurs et aux robots d'interpréter les mouvements de la tête ou des yeux d'un utilisateur comme une commande (déplacement du curseur, clic, etc.). Des applications dans le domaine de la santé, de la recherche et de la robotique à l'aide de technologies utilisant des méthodes de suivi du regard brevetées sont en cours de réalisation.

STATUT

Vainqueur de l'Arkathon Hacking Health Valais Bourse « Coup de pouce », Fondation R. Liechti
Intégration au programme The Ark Fondation de la société Eyeware Tech SA à Martigny.

DIALEXY

ÉQUIPE

- Mónica Camino Moreno, entrepreneur et traductrice certifiée
- Blaise Potard, postdoctorant à l'Idiap et à CereProc Ltd (Edimbourg, GBR)
- Albert Solà, développeur web senior à Lyles-Sutherland (Edimbourg, GBR)

OBJECTIF

Développement d'algorithmes dans le domaine du machine learning en vue d'améliorer et automatiser les processus de traduction et d'obtenir une traduction de grande qualité.

STATUT

Fondation de la société Dialexy à Edimbourg (GBR).

APP4 AUTISM

ÉQUIPE

- Cristina Costescu, assistante de recherche, Department of clinical psychology and psychotherapy, Babeş-Bolyai University, Cluj-Napoca (Roumanie)
- Josef Kalleder, Apator SA Poland, directeur des ventes pour la Roumanie et la République de Moldavie
- Mihai Calinescu, coordinateur du département de production apprentissage sur vidéo et en ligne, éditeur vidéo, SC Frontline Solutions SRL, Cluj-Napoca (Roumanie)

OBJECTIF

Développement d'un instrument thérapeutique et éducatif moderne afin de faciliter la communication entre des enfants autistes et leurs familles et éducateurs.

STATUT

Une start-up est en voie de création en Roumanie.



Bon nombre de projets récompensés lors des précédentes éditions de l'ICC poursuivent leur aventure, avec parfois de beaux succès à la clé.

32

HORUS

Le projet d'assistant visuel pour les personnes malvoyantes a débouché sur la création de l'entreprise Horus Technology en Italie. Après avoir reçu de prestigieux prix (Idea Challenge d'EIT Digital, Unicredit Startlab et IBM Smartcamp), elle vient de recevoir un investissement de 900'000 USD par la holding américaine 5Lion.

ANEMOMIND

La plate-forme d'analyse et d'optimisation des performances des navires à voile a donné le jour à une entreprise homonyme sise à Ecublens (VD). Outre des succès sportifs, elle a été nommée au Design Award METS, lors du Marine Equipment Trade Show 2015 à Amsterdam et a reçu un financement de démarrage de 100'000 CHF de la part de la Fondation pour l'innovation technologique.

BIOWATCH*

La montre biométrique se basant sur le réseau de veines au niveau du poignet a remporté le prix des employés UBS à l'occasion du concours mondial de l'innovation UBS Future of Finance Challenge 2015. Grâce à un financement de la Commission pour l'innovation et la technologie (CTI), la start-up poursuit son aventure en collaborant avec le Centre suisse d'électronique et de microtechnique (CSEM) à Neuchâtel et l'Idiap.

REMEETING*

Enregistrer et archiver des réunions, puis pouvoir effectuer des recherches efficaces dans leur contenu? Une idée qui a valu à ReMeeting, devenue recapp IT AG* entretemps, de recevoir de prestigieuses récompenses telles que le Grand Prize de UC Berkeley Startup Competition (LAUNCH). La jeune pousse a connu le soutien de la Fondation Hasler pour transformer son innovation technique en produit commercialisable.

LUMA7

La technologie d'analyse de texte de cette jeune entreprise valaisanne, grâce à laquelle il est possible de faire des recherches dans le contenu et de comprendre les relations entre les différents éléments, est utilisée par des entreprises de médias suisses ainsi que le parlement du Valais.

CLOUDGUIDE

L'entreprise a créé une application permettant d'utiliser son smartphone comme guide audio lors de la visite d'un musée. Cloudguide propose les outils nécessaires aux musées pour créer leur propre contenu et le mettre à disposition de la clientèle sur l'application mobile de la start-up. Cette idée a déjà séduit de nombreuses institutions de par le monde. L'entreprise a également reçu le Prix de l'Innovation de la banque espagnole Caja de Ingenieros.

*Spin-off de l'Idiap

Un système de reconnaissance vocale multilingue pour la restauration rapide

Dans le cadre d'un projet financé par la Commission pour la technologie et l'innovation (CTI), l'Idiap et l'entreprise suisse Veovox ont planché sur le développement d'un système de reconnaissance vocal multilingue particulièrement adapté aux conditions bruyantes pouvant régner dans un restaurant.

Lorsque le client d'un restaurant drive-in effectue la commande d'un repas, un opérateur la transmet aux cuisines, où plusieurs employés vont la traiter avant de la remettre au consommateur. L'employé utilise ses deux mains pour préparer la commande et doit interrompre ses gestes pour valider les articles ajoutés à l'aide d'une interface tactile. Les erreurs sont aujourd'hui relativement fréquentes car le mécanisme de validation est mal adapté à ces conditions.

Réduire le nombre d'erreurs

Afin de pallier ce problème, l'entreprise Veovox a développé, en collaboration avec l'Idiap, un système basé sur la reconnaissance de la parole pour la gestion de commandes. Au fur et à mesure qu'ils remplissent le sachet, les employés nomment les articles qu'ils y déposent. Le système vérifie qu'ils correspondent bien à la commande. Si tel est le cas, ils sont effacés de la liste. Dans le secteur de la restauration rapide, la solution proposée par Veovox permet ainsi de réduire drastiquement le nombre d'erreurs de préparation de commande, à la plus grande satisfaction des clients. Véritable pièce maîtresse de cette technologie, le système de reconnaissance vocale a été développé à l'Idiap, avec le soutien du Dr Mathew Magimai Doss.

Un peu d'Idiap dans les hamburgers

Cette solution novatrice a été développée entre 2013 et 2015 dans le cadre du projet MultiVEO et soutenue par la CTI. Ce projet est la continuation logique de deux précédentes collaborations fructueuses entre les partenaires. Dès le début du projet, des tests ont mis en concurrence l'approche de reconnaissance vocale développée par l'Idiap et Veovox avec d'autres solutions commerciales disponibles sur le marché. Les adaptations au scénario très contraint d'une liste limitée de commandes vocales ont permis à Veovox de proposer le système le plus compétitif. Plus important, l'approche utilisée par l'Idiap permet d'adapter le système rapidement à de nouvelles langues, sans devoir recourir à l'achat de données supplémentaires.

La prochaine fois que vous commanderez un hamburger, c'est peut-être un peu l'Idiap qui vous fera passer à table !

Transfert de technologie : faire fructifier les travaux de recherche

Dr Florent Monay et Dr Hugues Salamin ont repris en 2015 la gestion du bureau de transfert technologique (TTO, pour Technology Transfer Office) de l'Idiap. Anciens chercheurs devenus développeurs, ils ont à cœur de donner une seconde vie aux travaux de leurs collègues en faisant le lien avec des entreprises soucieuses d'innover.

Dans le domaine de la recherche fondamentale, le succès se mesure au nombre de publications et de citations qu'elles génèrent. Cette logique du « publish or perish » fait qu'en termes de carrière, un chercheur n'a pas forcément intérêt à consacrer du temps au développement d'un produit commercialisable.

Principalement financé par de l'argent publique, donc par les contribuables – au nombre desquels comptent les entreprises – l'Idiap a fait le pari du transfert de technologie depuis plusieurs années en se dotant d'un office dédié à cette tâche (TTO) et en soutenant une équipe de développeurs, aujourd'hui au nombre de douze. L'idée est d'offrir un retour sur investissement à l'industrie suisse et valaisanne.

« Ce sont les chercheurs qui font le premier pas, à nous ensuite de faire fructifier la prise de contact. »

Florent Monay et Hugues Salamin

Au-delà de la recherche

Depuis le début 2015, Florent Monay et Hugues Salamin ont repris les commandes du bureau de transfert de technologies. En parallèle, ils travaillent également comme développeurs. Issus du monde académique et armés du titre de docteur, leur profil est idéalement placé entre la recherche, le développement et l'application. « Le but de notre travail est de mettre en relation les chercheurs et les entrepreneurs. Pour cela il est essentiel de bien comprendre ces deux mondes, leurs besoins et leurs capacités », explique Florent Monay.

À leur arrivée au TTO, Hugues Salamin et Florent Monay ont commencé par mettre à jour le portfolio technologique de l'Idiap. Ils ont donc fait l'inventaire des technologies développées par les chercheurs ainsi qu'un état des lieux quant à leur degré d'aboutissement.

« Le portfolio comporte aujourd'hui une quarantaine de technologies, poursuit Hugues Salamin. Mais il est rare que l'une d'elles corresponde exactement au besoin précis d'une entreprise. Il nous sert donc principalement de vitrine, permettant ainsi de montrer à des partenaires potentiels des exemples concrets de ce que l'Idiap peut offrir. »

Lien entre recherche et industrie

Sur cette base, une entreprise peut ainsi décider de monter un projet et demander le soutien de la Commission pour l'innovation et la technologie (CTI) ou de la Fondation pour l'innovation en Valais The Ark. « À ce stade, reprend Florent Monay, notre travail consiste à trouver, au sein de l'Idiap, le chercheur le plus habilité à prendre la responsabilité scientifique. » C'est alors que les deux hommes aident les partenaires à rédiger la demande de financement.

En cas de succès, ils les assistent dans les démarches en vue de la dépose d'un brevet. « Cela se fait suite à la demande d'un chercheur. Elle passe devant le « patent committee », composé de la direction de l'Idiap, d'un membre du TTO et d'un chercheur qui a une longue expérience de brevets, explique Florent Monay. Le dépôt d'un brevet est une démarche longue et coûteuse. Nous cherchons donc à nous assurer de la valeur intrinsèque d'une invention avant d'entreprendre cette démarche. »

Motiver les chercheurs

Au niveau interne, la mise à jour du portfolio technologique vise également à motiver les chercheurs à être actifs dans le domaine du transfert de technologies. « Le portfolio et le transfert technologique permettent aussi d'identifier de nouveaux problèmes pour appliquer leur recherche et de bénéficier ainsi de ressources supplémentaires pour améliorer leurs algorithmes », poursuit Hugues Salamin.

Au terme de leur première année, les deux responsables du TTO en tirent un bilan positif. « En 2015, les chercheurs de l'Idiap ont demandé et obtenu le financement pour trois projets The Ark et quatre projets CTI, et ont déposé trois demandes de brevet dans les domaines de la biométrie et du traitement de la parole », commentent-ils.

En conclusion, Hugues Salamin et Florent Monay se réjouissent de ce que les entreprises connaissent l'Idiap : « Ce sont les chercheurs qui font le premier pas, lors de séminaires, salons ou conférences. A nous ensuite de faire fructifier la prise de contact et de fidéliser les entreprises qui ont placé leur confiance dans l'institut ! »





visages
36 — 43

La statistique comme langage interdisciplinaire

Spécialiste en mathématiques appliquées et en statistiques, David Ginsbourger cultive le goût de l'interdisciplinarité. Sa trajectoire l'a ainsi amené à des collaborations dans des domaines aussi variés que l'ingénierie, l'hydrogéologie ou les sciences du climat.

Lorsque l'on entend un chercheur Idiap parler de « machine learning », on s'attend à le voir évoquer des techniques de reconnaissance visuelle ou de robotique. Aussi peut-on être surpris quand David Ginsbourger l'associe au climat. Mais, outre son intérêt pour la statistique, c'est un goût prononcé pour les interactions interdisciplinaires qui ressort.

Pour ce chercheur, « des approches similaires apparaissent en parallèle dans différentes disciplines, avec des usages et des terminologies pouvant varier significativement d'un champ à l'autre. Le recours au langage mathématique et à des concepts statistiques permet souvent d'y voir plus clair. Les chercheurs en machine learning viennent d'horizons variés, avec leurs propres codes, modes de pensée et buts, mais ils se retrouvent sur bon nombre de questions de fond. »

Expériences in silico

Après un double diplôme à l'École des mines de Saint-Etienne et à la Technische Universität de Berlin, David Ginsbourger entame à Saint-Etienne un travail de doctorat en mathématiques appliquées. Il travaille sur des problématiques de planification d'expériences numériques. « La simulation est employée dans de nombreux domaines pour compléter voire se substituer à des expériences réelles. Mais les expériences numériques elles-mêmes ont un coût, notamment en temps de calcul, si bien que

« Des approches similaires apparaissent en parallèle dans différentes disciplines. »

David Ginsbourger
Chercheur senior



des stratégies de planification d'expériences sont utilisées. » Lors d'une visite aux Etats-Unis, David Ginsbourger croise le parcours de l'hydrogéologue Philippe Renard. Il rejoint plus tard son équipe à l'Université de Neuchâtel, en tant que collaborateur scientifique.

L'Idiap prend pied en Suisse alémanique

La vie d'équipe bat son plein lorsque David Ginsbourger se voit proposer un poste de maître-assistant à l'Institut de statistique mathématique et d'actuariat de l'Université de Berne. « J'ai débarqué dans un institut possédant à la fois une tradition de haut niveau théorique et de liberté académique, ce qui m'a permis d'approfondir mes connaissances statistiques tout en cultivant mes collaborations interdisciplinaires », explique-t-il.

Après son habilitation, il cherche un poste de professeur ou de chercheur permanent lorsque se présente l'occasion

d'intégrer l'Idiap. « Nous avons réussi à mettre sur pied un poste à cheval sur les deux sites », poursuit le scientifique. Depuis septembre 2015, il travaille au développement d'un groupe de recherche sur le thème de la quantification des incertitudes et de l'optimisation, tout en gardant un engagement à temps partiel à Berne. Cette nomination arrive au bon moment dans la carrière de David Ginsbourger : « Elle

va me permettre de me lancer dans un nouveau cycle de projets de recherche originaux dans le contexte très dynamisant de l'Idiap ». Fidèle à son goût pour l'interdisciplinarité, il se réjouit de s'y lancer en interaction avec les autres chercheurs permanents de l'Idiap.

La face cachée des projets de recherche



38

Dans le monde de la recherche comme ailleurs, l'argent est bien souvent le nerf de la guerre. À l'Idiap, Elisa Bovio accompagne les chercheurs dans leurs demandes de financement et tient la comptabilité des projets de recherche.

Après avoir travaillé pour une compagnie d'assurance, Elisa Bovio, jeune diplômée en économie d'entreprise de la Haute école de gestion de Sierre, a rejoint en juin 2015 les rangs de l'administration de l'Idiap. Elle y a une double casquette d'aide-comptable et de Junior Program Manager.

Aider les scientifiques

En sa qualité d'aide-comptable, elle s'occupe des salaires, mais aussi du remboursement des notes de frais. Elle se charge également de la comptabilité analytique des projets de recherche, veillant ainsi au bon respect des divers postes budgétaires. Cette dernière tâche lui permet de mieux faire le lien avec sa deuxième mission de gestion administrative de projets.

En tant que Junior Program Manager, Elisa Bovio aide avant tout les chercheurs dans le travail de soumission de projets et gère les contacts avec les nombreux partenaires. Par exemple, elle assiste le Dr Jean-Marc Odohez dans le projet européen MuMMER (voir p. 26) et gère des projets financés par l'agence américaine DARPA ainsi que des subsides de la fondation Hasler et du programme européen H2020. Un travail très diversifié, ouvert vers l'intérieur et l'extérieur, et particulièrement motivant.

Un environnement cosmopolite

L'institut Idiap, Elisa Bovio l'avait croisé par hasard durant son bachelor, à l'occasion d'une visite organisée pour la HES. « À l'époque, dit-elle, je ne m'imaginai pas travailler dans le domaine de la recherche, qui me paraissait bien lointain. » Se trouvant un peu à l'étroit dans le monde de l'assurance, cette passionnée de yoga et de voyages décide de changer de cap pour intégrer les rangs de l'Idiap. « J'y ai trouvé un environnement et des gens passionnants. J'apprécie tout particulièrement le côté multilingue de l'institut et la diversité des parcours, qu'ils soient personnels ou professionnels. J'ai l'âme vagabonde — j'ai déjà passé près d'une année en Amérique du Nord et voyagé plusieurs fois en Asie du Sud-Est — j'apprécie d'autant plus cette ambiance cosmopolite. »

« J'y ai trouvé un environnement et des gens passionnants. »

Elisa Bovio

Aide-comptable et Junior Program Manager

Elisa Bovio aime également la variété des tâches qui lui sont confiées, et la liberté qui va avec. « Avoir deux casquettes est à ce niveau un vrai plus. Mon emploi du temps varie beaucoup sur l'année, cela évite la routine et permet de jongler au mieux avec les besoins des chercheurs et de l'institut. Cette année, j'ai également pu m'investir dans

l'organisation du séminaire ISO sur les normes en biométrie ainsi que dans les activités liées au bicentenaire de l'entrée du Valais dans la Confédération. L'année prochaine, je reprendrai le dossier de l'International Create Challenge (voir pp. 31-32). Ce sera encore une fois l'occasion de plonger au cœur des nouvelles technologies, un domaine dont je suis de plus en plus friande ! »

Suivez mon regard !



Sur la lancée de l'ICC'2015, qu'il a brillamment remporté, Kenneth Funes Mora décide d'exploiter les résultats de son travail de doctorat en créant sa propre entreprise, Eyeware Tech. Parallèlement, ce spécialiste du suivi du regard en 3D poursuit sa carrière de chercheur à l'Idiap en tant que postdoc.

C'est un long chemin et une bourse Erasmus Mundus qui ont mené ce chercheur du Costa Rica jusqu'à Martigny ! Tout débute avec une licence obtenue à l'Instituto Tecnológico du Costa Rica, suivie d'un passage par les universités Heriot-Watt d'Edimburgh, Dijon, Gérone, puis à l'INRIA de Grenoble, où il effectue son travail de master.

Il rejoint l'Idiap et le groupe du Dr Jean-Marc Odobez en janvier 2011, et se lance dans un doctorat portant sur le suivi du regard en 3D. Ces travaux ont de multiples débouchés, tant au niveau de la recherche qu'à celui des applications. Le but est de pouvoir déterminer à tout moment où sont dirigés les yeux et l'attention d'une personne sans interfaces particulièrement sophistiquées. Des chercheurs en sciences humaines peuvent s'en servir pour étudier des interactions sociales, mais cette technologie a également des applications dans les domaines de la robotique, des sciences cognitives ou encore du marketing.

L'innovation, une aventure !

Au moment de terminer sa thèse — il décroche le titre académique en octobre 2015 —, il se lance dans l'aventure de l'innovation en s'inscrivant à l'Arkathon, le premier hackaton suisse dédié à la santé, organisé par The Ark. Avec son ami et collègue Carlos Becker, lui-même en train de réaliser un doctorat à l'EPFL dans le domaine de la Computer Vision, il planche sur une application de ses travaux de recherche. « Le visage humain est très expressif, explique Kenneth Funes Mora, et nos algorithmes de reconnaissance sont à même de le déchiffrer. En utilisant des expressions prédéterminées comme code, nous pourrions proposer à des personnes tétraplégiques de piloter facilement un ordinateur en associant des mouvements de la tête et du visage à des commandes. Il existe déjà de tels systèmes mains libres, mais ils sont coûteux et nécessitent souvent du matériel spécifique (des points réfléchissants à poser sur le front ou sur un casque) que seule une personne extérieure peut installer. Notre système, qu'on pourrait décrire comme une souris virtuelle, ne requiert quant à lui qu'une simple caméra 3D, un outil appelé à remplacer bientôt la webcam qui équipe la plupart des ordinateurs. »

Suite au succès rencontré lors de l'Arkathon, le duo décide de s'inscrire à l'ICC'2015 (voir article pp. 31-32). La compagnie de notre chercheur / entrepreneur le met alors en relation avec Serban Mogos, une connaissance au bénéfice d'un doctorat en Entrepreneurship de la Carnegie Mellon University. Les compétences s'additionnant, le projet de souris virtuelle se développe et prend forme. Si bien que le trio remporte le concours. « Ces journées étaient longues et intenses, mais nous y avons appris beaucoup. C'était très stimulant, vraiment une très belle expérience », explique Kenneth Funes Mora.

Après avoir reçu une bourse « Coup de pouce » de la Fondation Dr René Liechti, le trio fait le grand saut. En compagnie de Jean-Marc Odobez, ils lancent leur propre entreprise, Eyeware Tech SA, avec le soutien de l'Idiap ainsi que de l'incubateur IdeArk.

« Nous aimerions mettre notre système de souris virtuelle sur le marché et l'améliorer en y intégrant la technologie de suivi du regard en 3D sur laquelle j'ai travaillé durant ma thèse. Une tâche que je vais poursuivre sous la forme d'un postdoc à l'Idiap, toujours dans le groupe de Jean-Marc Odobez. Je me réjouis de poursuivre ma collaboration avec Jean-Marc, lui qui a su m'accompagner durant mon doctorat et la naissance d'Eyeware Tech. »

25 bougies pour l'Idiap

2016 marque la 25^e année d'activité de l'Idiap. De sa fondation à son jubilé, retour sur 25 événements qui ont fait l'histoire de cet institut aujourd'hui devenu adulte.

Timeline de l'Idiap

1991

Avec le soutien de la Fondation Dalle Molle pour la qualité de la vie, l'Idiap voit le jour et s'installe dans la Villa Tissières, à Martigny.

1996

Engagement d'un nouveau directeur en la personne du Professeur Hervé Bourlard.



2001

Début du Pôle de recherche national « Gestion interactive et multimodale de systèmes d'information » (IM2), dont l'Idiap assure la coordination.

2002

L'Idiap fait le buzz en analysant des enregistrements d'Oussama Ben Laden.

2002

L'Idiap s'agrandit et construit le pavillon Dalle Molle en ville de Martigny.

2002

La 12^e conférence IEEE (Institute of Electrical and Electronic Engineers) en « Neural Networks for Signal Processing » est la première conférence internationale organisée par l'Idiap à Martigny.

2003

Premier projet avec les Etats-Unis.

2004

Financement d'une nouvelle volée de projets européens de grande envergure (AMI/AMIDA) à fort impact international.

2005

Création de IdeArk SA, incubateur de start-up ayant pour mission de valoriser les technologies de l'Idiap.

2007

Lancement de Klewel, première spin-off issue de l'Idiap.



2007

Déménagement de l'Idiap dans les locaux du Centre du Parc, à Martigny.

2008

Signature de l'alliance stratégique avec le domaine des EPF.



2009

Ouverture d'un showroom dans les locaux de l'institut.

2010

L'Idiap met à disposition du marché la première version de son portefeuille de technologies.

2011

L'Idiap signe un accord de collaboration avec les prestigieux Institut indien de technologies de Guwahati (IITG) et Institut international des technologies de l'information de Hyderabad (IIIT).

2011

Premiers titres académiques EPFL décernés à trois chercheurs permanents.

2012

Les comptes d'exploitation de l'Idiap dépassent la barre mythique des 10 millions pour la première fois.

2012

Le projet MediaParl, soutenu par la Loterie Romande et le Grand Conseil valaisan, démarre.

2012

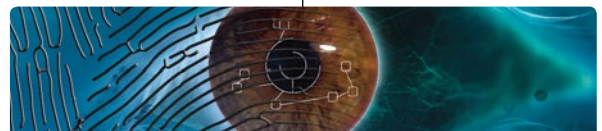
Création de l'Association Interactive Multimodal Information Management (AIM2), suite d'IM2.

2012

Lancement du premier International Create Challenge (ICC'2012).

2013

Validation des premiers brevets déposés par l'Idiap.



2014

Lancement du « Swiss Center for Biometrics Research and Testing », premier centre européen en biométrie.

2014

Sous l'impulsion de son directeur, sept personnalités du monde scientifique mènent un audit approfondi des activités de l'Idiap.

2015

L'Idiap participe aux festivités du bicentenaire de l'entrée du Valais dans la Confédération avec son projet Valais*Wallis Digital.

2016

Pour son 25^e anniversaire, l'Idiap choisit d'organiser l'« Idiap Innovation Day », pour présenter ses activités aux entreprises romandes.

Prix Idiap

Interne

Chaque année, l'Idiap décerne deux prix destinés à ses doctorants. Le premier récompense une recherche, le second une publication. Pour l'attribution du prix Idiap de la Recherche, le candidat est évalué par une commission interne sur la base de cinq critères : ses publications, sa collaboration dans l'équipe, son implication dans le projet, son sens de la communication et son autonomie. Pour le prix de la Publication, une première sélection est effectuée par les seniors 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 2015, le prix de la Recherche a été attribué à **Ivana Chingovska** pour l'excellence de ses recherches traitant de l'antispoofing en biométrie, appliqué à la reconnaissance faciale, pour la quantité et qualité de ses publications, spécialement en 2015, et pour sa participation à la communauté des doctorants Idiap et EDEE/EPFL. Celui de la publication a été remis à **Nikolaos Pappas** pour son excellent article scientifique intitulé « Adaptive sentiment-aware one-class collaborative filtering ».



Nikolaos Pappas



Ivana Chingovska

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 2015, WURZEN, SEPTEMBER 2015,
 WINNER OF THE PERSON DISCOVERY SUBTASK
EUMSSI team at the MediaEval Person Discovery Challenge

M. RAZAVI, R. RASIPURAM, M. MAGIMAI DOSS
 LANGUAGE AND TECHNOLOGY CONFERENCE (LTC)
 2015, POZNAN, POLAND, BEST STUDENT PAPER AWARD
Pronunciation Lexicon Development for Under-Resourced Languages Using Automatically Derived Subword Units : A Case Study on Scottish Gaelic

C. MCCOOL, R. WALLACE, M. MCLAREN,
 L. EL SHAFEY, S. MARCEL
 IET 2015 BIOMETRICS PREMIUM BEST PAPER AWARD
Session variability modelling for face authentication

N. CHACKO, K. CHAN ET M. LIEBLING
 IEEE INTERNATIONAL SYMPOSIUM ON BIOMEDICAL
 IMAGING, BROOKLYN, NY, APRIL 2015, BEST STUDENT
 PAPER AWARD
Intensity-Based Point-Spread-Function-Aware Registration for Multi-View Applications in Optical Microscopy

Thèses achevées

Sept étudiants ont achevé leur thèse en 2015 : Kenneth Alberto Funes Mora, Ashtosh Sapru, Laurent Son Nguyen, Mohammad J. Taghizadeh, Maryam Habibi, Sree Harsha Yella et Ivana Chingovska.

3D GAZE ESTIMATION FROM REMOTE RGB-D SENSORS

Kenneth Alberto Funes Mora

29 mai 2015

Directeur de thèse : D^r Jean-Marc Odobez

Membres du jury : Prof. Kamiar Aminian,
D^r Ronan Boulic, Prof. Dan Witzner Hansen
et Prof. Louis-Philippe Morency

AUTOMATIC SOCIAL ROLE RECOGNITION AND
ITS APPLICATION IN STRUCTURING MULTIPARTY
INTERACTIONS

Ashtosh Sapru

28 avril 2015

Directeur de thèse : Prof. Hervé Bourlard

Membres du jury : Prof. David Atienza Alonso,
Prof. Fabio Pianesi, D^r Alessandro Vinciarelli,
D^r Jean-Marc Vesin

COMPUTATIONAL ANALYSIS OF BEHAVIOR
IN EMPLOYMENT INTERVIEWS AND VIDEO RESUMES

Laurent Son Nguyen

18 février 2015

Directeur de thèse : Daniel Gatica-Perez

Membres du jury : Prof. Jean-Philippe Thiran,
Prof. Marianne Schmid Mast, Prof. Mohamed
Chetouani, Prof. Sabine Sússtrunk

ENABLING SPEECH APPLICATIONS USING AD HOC
MICROPHONE ARRAYS

Mohammad Javad Taghizadeh

27 mars 2015

Directeurs de thèse : Prof. Hervé Bourlard
et D^r Philip N. Garner

Membres du jury : Prof. Jean-Philippe Thiran,
Prof. Emanuel Habets, Prof. Dietrich Klakow,
D^r Hervé Lissek

MODELING USERS' INFORMATION NEEDS
IN A DOCUMENT RECOMMENDER FOR MEETINGS

Maryam Habibi

30 novembre 2015

Directeurs de thèse : Prof. Hervé Bourlard
et D^r Andrei Popescu-Belis

Membres du jury : Prof. Jean-Philippe Thiran,
D^r Jim Glass, Prof. Stéphane Marchand-Maillet,
D^r Martin Rajman

SPEAKER DIARIZATION OF SPONTANEOUS MEETING
ROOM CONVERSATIONS

Sree Harsha Yella

19 janvier 2015

Directeur de thèse : Prof. Hervé Bourlard

Membres du jury : Prof. Anja Skrivervik,
D^r Andreas Stolcke, D^r Xavier Anguera,
D^r Jean-Marc Vesin

TRUSTWORTHY BIOMETRIC VERIFICATION
UNDER SPOOFFING ATTACKS :

APPLICATION TO THE FACE MODE

Ivana Chingovska

23 novembre 2015

Directeurs de thèse : Prof. Hervé Bourlard
et D^r Sébastien Marcel

Membres du jury : Prof. Sabine Sússtrunk,
Prof. Julian Fierrez-Aguilar, Prof. Raymond Veldhuis,
Prof. Jean-Philippe Thiran



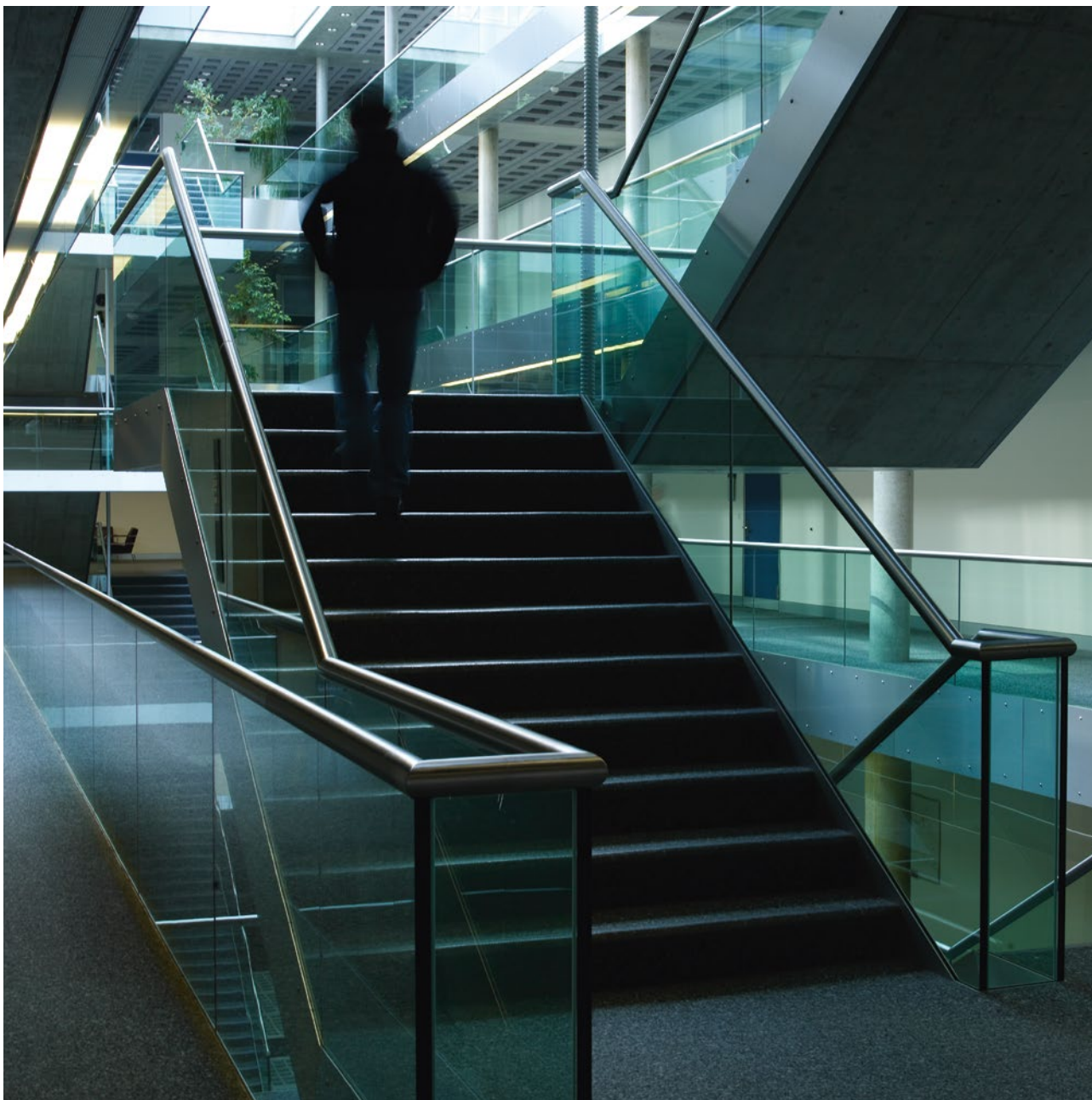
Centre du Parc, Rue Marconi 19,
case postale 592, CH - 1920 Martigny

T +41 27 721 77 11

F +41 27 721 77 12

M info@idiap.ch

www.idiap.ch



scientific inserts

I — XLIV

Research areas

To face its continuous growth and diversification in different, but complementary, research directions, while building upon its general institutional theme of "Human and Media Computing" (as defining ourselves on our website www.idiap.ch),

Idiap's expertise mainly spans five research areas, which are summarized in Figure 1 below. In that table, each research area is also described in few keywords. As described later, those generic research areas are covered by 10 research group.

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

Figure 1

Overview of Idiap research areas.

PERCEPTUAL AND COGNITIVE SYSTEMS

Speech processing; Natural language understanding and translation; Document and text processing; Vision and scene analysis; Multimodal processing; Cognitive sciences and Robotics.

Idiap combines its multi-disciplinary expertise to advance the understanding of human perceptual and cognitive systems, engaging in research on multiple aspects of human-computer interaction with computational artefacts such as natural language understanding and translation, document and text processing, vision and scene analysis (with a particular emphasis on human sensing), multimodal interaction, computational cognitive systems, robotics, and methods for automatically training such systems.

HUMAN AND SOCIAL BEHAVIOUR

Social media; Mobile media; Social interaction analysis; Social signal processing.

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

INFORMATION AND PRESENTATION INTERFACES

Multimedia information systems, User interfaces; System evaluation.

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

BIOMETRIC PERSON RECOGNITION

Face recognition (detection-localization-identification-verification); Speaker identification and verification; Multimodal biometric person recognition; Counter-measures to spoofing attacks.

Biometric person recognition (Biometrics) refers to the process of automatically recognizing a person using distinguishing behavioural patterns (gait, signature, keyboard typing, lip movement, hand-grip) or physiological traits (face, voice, iris, fingerprint, hand geometry, EEG, ECG, ear shape, body odour, body salinity, vascular). Idiap is currently a recognized leader in that field, fully exploiting our multi-disciplinary expertise in image processing, computer vision, pattern recognition and machine learning. This expertise has recently been confirmed by the official launch of a 'Swiss Center for Biometric Security Research and Testing' (<http://www.biometrics-center.ch/>) hosted at Idiap.

MACHINE LEARNING

Statistical and neural network based machine learning; Computational efficiency, targeting real-time applications; Very large datasets; Online learning.

Research in machine learning aims at developing computer programs able to learn from examples. Instead of relying on a careful tuning of parameters by human experts, machine-learning techniques use statistical methods to directly estimate the optimal setting, which can hence have a complexity beyond what is achievable by human experts. Today, Idiap is also recognized as a key leader in that field with new trends towards "collaborative machine learning", deep neural network architectures, and large-scale distributed learning algorithms.

NEW RESEARCH THEMES INITIATED IN 2015

In addition to these "horizontal" research themes, new activities recently started to be developed towards bio-medical applications (through the new Computational Biomedicine Group) and Environmental Modeling (through the new group in "Uncertainty Quantification and Optimal Design").

1. Speech and Audio Processing

OVERVIEW

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

GROUP OVERVIEW

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

The Speech and Audio Processing group in 2015 was composed of 1 head of group, 3 Principal Investigators, 9 postdocs, 1 scientific collaborator, 10 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

We have always used both neural network and Markov model based approaches, leading to the KL (Kullback Leibler)

HMM approach to fusion of neural network and HMM based speech recognition, especially for multilingual speech modelling. Use of techniques from HMM based speech recognition in HMM based speech synthesis resulted in a unified approach to speech recognition and synthesis. The group was well placed to take advantage of recent advances in "deep" neural networks (DNNs) technology, and the associated Kaldi automatic speech recognition (ASR) toolkit. Template-based speech processing has been always in attention to utilize the power of big data with less statistical assumptions. In this context, Idiap is recognized to exploit neural network posterior features as speech templates, and structured sparsity of posterior features have led to hierarchical frameworks for ASR, query-by-example spoken term detection (QbE-STD), speech coding and linguistic parsing applications. Several key contributions in microphone arrays, including ad-hoc microphone arrays, are well recognized. Sparse model based approaches have led to novel algorithms for source separation. The agglomerative clustering approach to speaker diarization is well regarded. Finally, the work in this group also resulted in several spin-offs, including dev-audio, Koemei and, more recently (end of 2014), recapp.

Much of the group's recognition has come from contributions to the speech and audio community; these include the AMI corpus¹ and the Juicer ASR decoder². More recently, the MediaParl³ and Walliserdeutsch⁴ databases have been released; other contributions such as ISS and SSP are on the GitHub site⁵. This highlights that, although the group has traditionally worked with English speech, many recent projects are multi-lingual in nature. This has in turn become a focus of the group; one on which we are well placed to capitalise given our geographical location. The group is also very active in contributing to the Kaldi open-source ASR toolkit⁶.

AUTOMATIC SPEECH RECOGNITION

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

- **Robust parametrization and acoustic modeling**

We are still investigating new features (e.g., posterior-

¹ www.idiap.ch/mmm/corpora/ami

² juicer.amiproject.org/juicer/

³ www.idiap.ch/dataset/mediaparl

⁴ www.idiap.ch/dataset/walliserdeutsch

⁵ github.com/idiap

⁶ kaldi.sourceforge.net

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

⁸ www.idiap.ch/scientific-research/projects/a-muse

⁸ www.idiap.ch/scientific-research/projects/a-muse

⁹ www.idiap.ch/scientific-research/projects/phaser

⁹ www.idiap.ch/scientific-research/projects/phaser

⁹ www.idiap.ch/scientific-research/projects/phaser

based features) and new acoustic models (new forms of hidden Markov models, such as KL-HMM, or artificial neural networks) that are more robust to noise and acoustic environments, as well as to speaker variability (e.g., accented speech). Over the last 12 months, our research has largely focussed on the use of DNNs for far-field speech recognition; novel approaches to adapt DNNs across different domains and languages; sparse modeling and low-rank representation to structure DNN testing posteriors according to the training domain for unsupervised acoustic model adaptation; development of articulatory feature (phonological feature) based ASR system; development of a unified framework for recognition and detection; (funded through the EC Eurostars project D-Box⁷ SNSF projects like A-MUSE⁸, PHASER⁹, FlexASR¹⁰, and IM2 NCCR sub-project Walliserdeutsch, HASLER Foundation project AddG2SU¹¹, CTI project ScoreL2¹², and Samsung Electronics Co., Ltd. (South Korea) sponsored project AMASSE¹³).

• **Cross-lingual and multi-lingual speech recognition**

Within the scope of the D-Box project, we focus on investigating and exploiting fast acoustic model adaptation techniques in cross-lingual and multi-lingual scenarios. The resulting speech recogniser relies on a hybrid approach, where an artificial neural network acoustic model is bootstrapped using well-resourced data and adapted to the target language (or environment). In the context of an industry sponsored project (Samsung, Inc. South Korea sponsored project DAUM¹⁴), we are also investigating Markov model based acoustic model adaptation techniques (currently referred to as "subspace Gaussian modeling") for different domains (e.g., accent or dialect).

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

• **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 (illustrated in Figure 4), 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. The subject of dialects remains pertinent as an important local focus, with plans underway to expand to other regions and countries such as Germany.

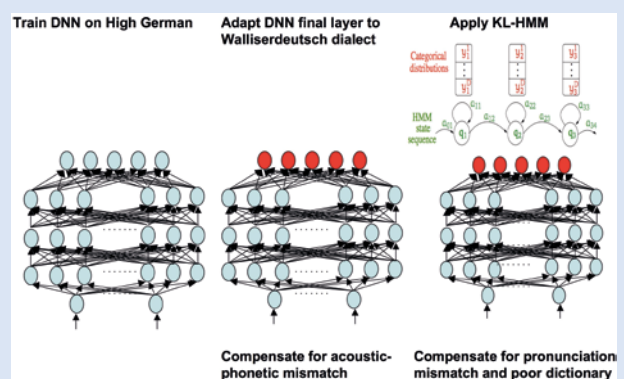


Figure 4

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

• **Template-based ASR**

Idiap continues to work on template-based or exemplar-based ASR using posterior-based features in the context of projects like SNSF project A-MUSE, SNSF project PHASER, and CTI project MultiVEO¹⁵. In particular, our current research is focussing on linking template-based speech recognition and statistical sequence model based speech recognition. Towards this end there are two on-going efforts. The first focuses on development of a "sparse" HMM framework by establishing a link between statistical speech recognition formalism and the recent theories of compressive sensing and sparse modeling. The framework integrates the advantages of template-based modelling within the HMM sequence matching scheme. The second effort focusses

¹⁰ <https://www.idiap.ch/scientific-research/projects/flexasr>

¹² <https://www.idiap.ch/scientific-research/projects/scorel2>

¹⁴ <http://www.idiap.ch/scientific-research/projects/daum>

¹¹ <https://www.idiap.ch/scientific-research/projects/addg2su>

¹³ <http://www.idiap.ch/scientific-research/projects/amasse>

¹⁵ <https://www.idiap.ch/scientific-research/projects/multiveo>

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. Over the past 12 months, this vision was realized through sparse modeling of posterior features. It has been shown that dictionary learning for sparse recovery of posteriors enables enhanced acoustic modeling for statistical ASR through hierarchical reconstruction of posterior templates (funded by SNSF project PHASER).

• Grapheme-based ASR

Over the past five 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 grapheme-to-phoneme relationship through the acoustic speech signal (funded through the SNSF-project FlexASR and HASLER Foundation project AddG2SU). In the past 12 months, our research activities in that direction have primarily focussed on exploiting the approach for lexical resource development through (a) acoustic data-driven grapheme-to-phoneme conversion and (b) derivation of "phone-like" subword units automatically from the acoustic data.

These research directions are particularly addressing practical challenges such as non-native or accented speech recognition, rapid development of ASR systems for new or under-resourced languages and/or domains. In addition, building on top of the core ASR research, we are also actively carrying out research on development of novel approaches for (a) utterance verification, (b) keyword spotting (spoken term detection), (c) objective assessment of coded speech and synthetic speech, and (d) automatic assessment of second language learners (funded through SNSF project A-MUSE, Idiap internship project, EC Eurostars project D-Box, armasuisse¹⁶, CTI project ScoreL2¹⁷). Keyword spotting is also addressed in the Applied Machine Learning group.

SPEECH SYNTHESIS

• Text-to-speech synthesis (TTS)

Although newer than ASR, TTS is now an established venture for the speech group at Idiap. TTS is central to two ongoing projects: SIWIS¹⁸ (Spoken Interaction With Interpretation in Switzerland), D-BOX, is DBOX still active? and SP2¹⁹ (SCOPES project on speech prosody).

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

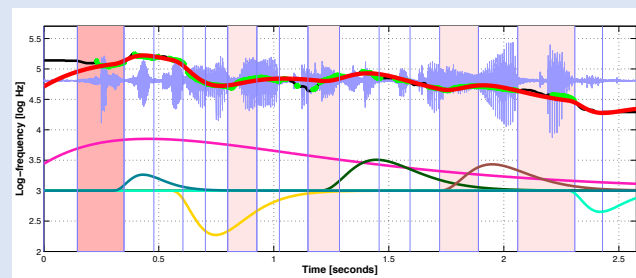


Figure 5

Modelling prosody using distinct atomic muscular twitches.

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 are developing novel models of speech prosody. The goal of SIWIS are mirrored in the SP2 project with a focus on eastern-European languages. Figure 5 illustrates one of our recent approaches to decompose the pitch contour (superimposed on top of the speech signal) into "atomic" patterns (displayed below the speech signal and pitch contour), which can be related to the muscular twitches at the glottal source. SIWIS and SP2 highlight the requirement for a collaborative approach: Speech to speech translation requires a large infrastructure that is difficult to maintain at a single site. To achieve this, SIWIS brings together partners at ETHZ and the Universities of Geneva and Edinburgh; SP2 adds partners from Budapest, Skopje and Novi Sad.

R&D work in TTS is also one of active areas for D-BOX project. Besides integrating the engine into an open-gaming platform through the implemented API supporting real-time multi-

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

¹⁷ www.idiap.ch/scientific-research/projects/scorel2

¹⁸ www.idiap.ch/project/siwis/

¹⁹ www.idiap.ch/scientific-research/projects/sp2

²⁰ www.siip.eu

player response, research in fast adaptation of TTS engine towards target speaker or different language is carried out.

- **Speech coding**

Another research area requiring TTS is speech coding, where very low bit rates can be achieved by using a speech recogniser followed by a synthesiser. Previously, under the RECOD project funded by armasuisse, the technology lends itself to operate at 200–300 bits per second. In 2015, work on speech coding continued under the SP2 project, mainly on aspects concerning prosody.

SPEAKER RECOGNITION AND SPEECH ANALYTICS

Starting from 2014, within the scope of the SIIP (EC FP7 project)²⁰, there has been a rejuvenation of core speaker recognition research in the Speech and Audio Processing group. In particular, as illustrated in Figure 6, Idiap is involved in developing a cutting edge suspect identification solution for identifying speakers captured in lawfully intercepted calls, including multiple types of social-media information streams. In addition to core speaker recognition, the technology will capitalise on fusion of other speech analytics algorithms such as gender, age, language, or accent identification, keyword and taxonomy spotting and voice cloning detection. Speaker recognition is also addressed in the Biometric Person Recognition group, as discussed in Section 9.



Figure 6

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

AUDIO PROCESSING

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

- **Query-by-example spoken term detection**

Idiap continues studies on the subject of keyword spotting with a focus on searching large audio archives using audio queries. The task is referred to as query-by-example spoken term detection (QbE-STD). The leading edge systems rely on template matching using posterior features as templates. Posterior features are sparse, and sparsity indicates their low-dimensional attributes. In 2015, our research focused on exploiting sparse modeling for characterizing the low-dimensional subspaces of query and background posteriors. QbE-STD is thus cast as a subspace detection problem (funded through SNSF project A-MUSE).

- **Sparse structured representations**

Idiap continues theoretical studies of sparse structured representations of audio signals, more specifically on model-based sparse component analysis, with different practical applications, such as source separation and localisation based on a small number of microphones, room acoustic modelling, and new beamforming techniques to improve recognition of speech recorded from farfield microphones. Exploiting structured sparsity pertained to speech perception and multipath propagation enables auditory and acoustic informed source separation and localization for far field ASR (supported by SNSF project PHASER).

- **Microphone arrays**

Idiap also works on audio recorded from microphone arrays, especially those organized in "ad-hoc" geometries (funded originally through the NCCR IM2, and more recently PHASER). Recently, we mainly focused on the automatic calibration of arbitrary microphone array geometries, fusing Euclidean constraints into generic mathematical concepts. This in turn enables the localization required for high quality data acquisition. In this context, we also consider environmental ambiguities such as reverberation and overlapping speech, thus requiring multi-source localization using ad-hoc microphones in reverberant rooms. Recently, we investigated the properties of time difference of arrival matrices specifying relative positioning of a source with respect to the pairs of microphones. Their low-rank property was exploited in an algebraic formulation to enable denoising and robust source localization using ad hoc microphones.

²¹ <http://www.idiap.ch/scientific-research/projects/rodi>

²³ <http://www.inevent-project.eu/>

²² <https://www.idiap.ch/scientific-research/projects/dimha>

²⁴ <http://www.idiap.ch/scientific-research/projects/sesame>

- **Speaker diarization**

In speaker diarization—the task of annotating temporal regions of audio recordings with labels indicating “who spoke when”—Idiap focuses on processing spontaneous overlapping speech through the use of multiple sources of information investigated in the context of various projects, including the SNSF project RODI²¹, SNSF project DIMHA²² and EU project inEvent²³. Over the last 12 months, our research has focussed on (a) linking and uniquely identifying speakers across large speech databases by making use of multiple modalities, (b) extraction of novel features in the frameworks of i-vectors and ANNs, and (c) scaling the KL-HMM approach, originally developed for ASR, to speaker diarization.

- **Audio indexing**

In the context of the Hasler Foundation project SESAME²⁴, Idiap aims at advancing the state-of-the-art in speech processing and spoken language understanding for automatically structuring data from spoken cultural heritage archives. Technical challenges include speaker segmentation and clustering, role recognition as well as topic/story segmentation and tracking over very heterogeneous. In 2014, our research focused on unsupervised topic segmentation using social role information.

- **Very low bit-rate speech coding**

Over the last three years, Idiap has become active in the domain of very low bit-rate speech coding through a collaboration with armasuisse, in the context of the RECOD project. We have proposed a novel architecture for low bit rate coding based on neural networks. This coder differs from previous attempts by replacing HMMs by deep neural networks, and by using the phonological speech representation instead of the phonetic one. In addition, the coder further lowers the operating latency. In 2015, structured sparsity of phonological representations was discovered that led to construction of a small size codebook of sparsity patterns for very low-bitrate speech coding (supported by SNSF project PHASER); a patent application is submitted.

- **Objective speech intelligibility assessment**

In the context of armasuisse project OMSI-2015_ ARMASUISSE²⁵ and CTI project ScoreL2, 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. Furthermore, the proposed approach can be applied to diverse speech assessment tasks, namely, (a) assessment of speech communication systems based on both traditional and more recent, very low bit-rate coding schemes in the presence of background noise; (b) synthetic speech intelligibility assessment; and (c) assessment of accentedness of non-native speech.

ROADMAP FOR CONVERSATIONAL INTERACTION TECHNOLOGIES

Over the past two years, Idiap was involved in an EC funded large scale roadmapping exercise in the area of “Conversational Interaction Technologies” – ROCKIT (Roadmap for Conversational Interaction Technologies; <http://rockit-project.eu/> – which resulted in a detailed strategic (and interactive) roadmap for research and innovation in the area of multilingual conversational interaction technologies (CIT), compiled from more than 1,000 inputs from the research and innovation stakeholders.

The ROCKIT consultation brought together all of the major European stakeholders in multilingual and cross-sectorial CIT across two years of discussion about how to develop a strong innovation community with a common outlook. Our consultation resulted in a detailed understanding of what Europe could achieve and what the future demand for multilingual products, services and research will be.

A key aspect of ROCKIT was also to organise a European research and innovation community in the area of conversational interaction technologies, integrating a wide-range of commercial organisations with application and use links to the area. ROCKIT will be structured around a set of sector-based clusters including mobile applications, healthcare, education, games, broadcast media, robotics, law enforcement, and security.

²⁴ www.idiap.ch/scientific-research/projects/sesame

²⁵ www.idiap.ch/scientific-research/projects/omsi-2015

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, France, 2000; Habilitation, University of Paris XIII, 2006; EPFL MER).

RESEARCH 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, for instance for object detection and scene analysis, tracking of people and biological structures, and image recognition in general.

Our current research follows two main axes. The first is the study of novel generic learning methods to cope with feature spaces of large dimension, or large training sets. The second is the development of new algorithms for scene understanding, which involves practical problems such as object detection in natural scenes, or pedestrian tracking with a multi-camera setup.

The Computer Vision and Learning group in 2015 was composed of one head of group, four PhD students and one intern, with strong interactions with the software development team.

KEY SCIENTIFIC OUTPUTS

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

MACHINE LEARNING

- **Joint feature selection**

To classify signals using very large number of measurements, or features, it is often desirable to apply a pre-processing step

during learning which selects a "good" sub-set of them. Many criteria have been developed for that task over the years, but while it is easy to measure how good a single feature is, these techniques often fail at taking into account the joint quality of a set of features.

We have developed new methods which relies on a joint models of feature responses, under which we explicitly measure the information content of the subset of features we are selecting. Our main contribution is a "Gaussian compromise", to approximate the entropy of a mixture of Gaussian densities, for which no analytical or tractable expression exist, and that we use to approximate the mutual information.

- **Efficient clustering**

The k-means algorithm is one of the cornerstone 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 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 to this work on exact methods, 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 are interested in speeding-up "hard sample" extraction. Most of the state-of-the-art detection algorithms are trained in a sequential manner, improving iteratively the detector by training it with samples miss-classified by its previous version.

Since the performance of detectors can be extremely good, finding a sufficient number of hard samples is computationally-intensive, and in all available methods, has a cost proportional to the number of images to visit.

Our approach consists in using the prior structure provided for any large collection of images, that is its recursive decomposition into homogeneous sub-families, to concentrate computation in a principled manner. We organize the set of images into a recursive tree consistent with the similarity between images, and adapt the Monte-Carlo Tree Search to recursively apply a bandit approach to the exploration-exploitation dilemma.

In practice, our approach alternate between looking into set of images and images it has poor information about, and images from which it has observed a high success rate in finding hard samples.

• **High-dimension 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, that is a method able to automatically figure out that a person leaving the field of view of a camera re-enters later the field of view of that camera or another one.

x

Such a problem boils down to metric learning, 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. In the application at hand, the challenges is to deal with changes in illumination, sensor calibration, and geometry between cameras.

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



Figure 7

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

• **Playground learning with pose-indexed features**

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

• **Multi-camera tracking**

We have a long-lasting collaboration with the CVLab at EPFL around the design of a robust multicamera tracking system, which now takes place in the context of the SNSF WILDTRACK project, in collaboration with ETHZ. Our most recent results consist of a new model and optimization scheme to handle properly the appearance / disappearance of "containees" (e.g. pedestrians) given the presence/absence of "containers" (e.g. cars).

Our system is composed of the Probabilistic Occupancy Map, and a flow-based convex multitarget tracker. The former estimates in individual time frames marginal probabilities of presence of people by minimizing the Kullback-Leibler divergence between the "true" posterior under a joint generative model, and a product law. The latter connects these detections properly by optimizing a flow in a spatio-temporal graph. Both components are extremely efficient computationally and process in real-time multiple camera streams to produce state-of-the-art results. Open-source implementations of both components are available under the GPL3 license.

• **Advertisement substitution in sport events**

In a collaboration with E-S. Concept. SA, we are developing a novel pipeline of algorithms to perform advertisement substitution on-the fly in broadcast video streams. The objective of this collaboration is to deliver a solution able to automatically and accurately localize perimeter advertisement during a sport event, and replace it with another one.

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

3. Social Computing

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 2015 was composed of one group head, one scientific collaborator, one SNSF Ambizione Research Fellow, three postdoctoral researchers, four PhD students, one visiting faculty, and one visiting student. Research lines investigated in 2015 included: social media and phone data analytics for cities, mobile crowdsourcing, ubiquitous face-to-face interaction analysis, and analysis of Maya hieroglyphic media collections.

KEY SCIENTIFIC OUTPUTS

Published studies on (1) social media analytics and mobile and online crowdsourcing to document and understand urban perception in cities; (2) multimodal analysis of socioaffective attributes of social video users (YouTube) integrating audio-visual perception and video crowdsourcing; (3) multimodal inference of social constructs in employment interviews; and (4) Maya hieroglyphic retrieval integrating shape and language models. 11 EPFL PhD students graduated since 2002.

SOCIAL MEDIA ANALYTICS AND MOBILE CROWDSOURCING TO UNDERSTAND PHENOMENA IN CITIES

Our work in this domain spans several research lines. First, we designed and implemented a mobile crowdsourcing methodology to engage youth in cities to collect multimedia items (geo-localized images and videos) that document urban

concerns like waste management, danger, or accessibility. Field studies have been implemented in multiple cities in Mexico (Figure 8). Furthermore, we are studying how urban perception can be crowdsourced and modeled to understand how sensitive citizens are with respect to such issues, and how different urban dimensions relate to each other. This work was done in the context of the SenseCityVity project²⁶ (Mobile Sensing, Urban Awareness, and Collective Action) funded by EPFL Center for Cooperation and Development (CODEV), and carried out with a multidisciplinary team from IPICYT, Mexico. This work will be expanded through a new project grant from Mexico's National Council of Science and Technology. This work has received national press coverage in Mexico^{27 28}.

In other work, through the SNSF Youth@Night project²⁹ (A Multidisciplinary Study of Young People's Going Out and Drinking Behaviors, in collaboration with Addiction Switzerland and the University of Zurich), and the MCSC project³⁰ (Mi Casa es su Casa, supported by the EPFL-UNIL Program on Cross-Disciplinary Research on Science and Society), we are investigating the use of mobile and on-line crowdsourcing and social media analytics (Twitter, Foursquare, Airbnb) to characterize urban phenomena at scale in cities both in Switzerland and other world regions³¹. This includes the characterization of urban areas according to patterns of human activity or the state of road conditions, and the perception of social ambiance in both popular urban places and personal spaces (Figure 8). Our work in this domain has been recently covered in the media^{32 33}.

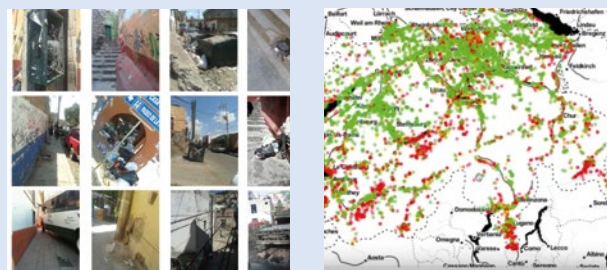


Figure 8

Left: SenseCityVity: Young people in Mexico document and reflect upon urban concerns. Right: Mapping social media activity in Switzerland from Twitter and Foursquare.

²⁶ <http://www.idiap.ch/project/sensecityvity>

²⁷ <http://www.conacytprensa.mx/index.php/centros-conacyt/2143-sensecityvity-estudio-de-la-geografia>

²⁸ <http://conacytprensa.mx/index.php/>

centros-conacyt/5198-crece-proyecto-sensecityvity-en-mexico-nota

²⁹ <https://www.youth-night.ch/>

³⁰ <http://www.idiap.ch/project/mcsc>

³¹ <https://www.youtube.com/>

[watch?v=UZT4bfVnL9Q](https://www.youtube.com/watch?v=UZT4bfVnL9Q)

³² <http://canal9.ch/comportements-humains-lidiap-sonde-les-smartphones/>

³³ <http://actu.epfl.ch/news/an-app-to-steer-clear-of-kenya-s-road-hazards/>

SOCIAL VIDEO ANALYTICS

This research line develops methods to characterize users of social video in sites like YouTube through the use of (1) video crowdsourcing techniques through platforms like Amazon Mechanical Turk to collect interpersonal judgments, which scales up to large amounts of data and demographically diverse viewers; and (2) automatic extraction of behavioral features including prosody, body motion, and facial expressions, which enables high-level social analysis. Our research, studying video resumes and video blogs, has shown that certain impressions of personality, mood, and hirability can be inferred to some degree from nonverbal and verbal cues. This research was supported until 2014 by the NISHA³⁴ project (NTT-Idiap Social Behavior Analysis Initiative) funded by NTT Communication Science Laboratories, Japan, and in 2015 has been done in the context of the SNSF UBImpressed project³⁵ (Ubiquitous Frist Impressions and Ubiquitous Awareness, Sinergia Interdisciplinary Program).

FACE-TO-FACE INTERACTION ANALYTICS

We are developing computational models to analyze dyadic and small-group interactions in workplace settings using a variety of sensors (cameras, Kinect devices, microphone arrays, smartphones, and wearables), and automatically infer key social variables for psychologists and recruiters (Figure 9). This is collaborative work with the University of Lausanne, Cornell University, and Vatel Switzerland. In the dyadic case, in the context of organizational behavior situations, we have investigated connections between hirability and personality. In the small-group case, we have studied links between behavioral cues and variables like emergent leadership and personality traits. In all cases, nonverbal cues are automatically measured from audio and video streams.

This research has been conducted in the framework of two SNSF projects, namely UBImpressed and SOBE, an Ambizione Research Fellowship. Our work in this domain has been covered by business media³⁶ and the European Commission as a H2020 Project Story³⁷.

VISUAL ANALYSIS OF MAYA HIEROGLYPHS

We are developing methods for automatic visual analysis of hieroglyphs for the three most celebrated ancient Maya codices, which are located in museums in Europe – Dresden, Paris, Madrid (Figure 10).



Figure 9
Sensors available in the social sensing lab. (1) Kinect V2, (2) Microcone, (3) Q-Sensor, (4) Android smartphone, (5) Google Glass.

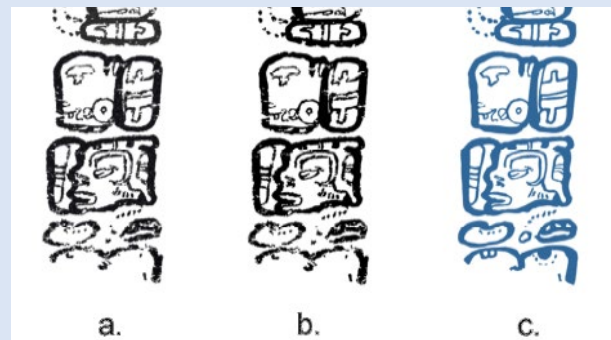


Figure 10
Preprocessing steps to generate Maya glyph data for visual analysis: (a) color images, (b) binarized images, (c) reconstructed glyphs by expert epigraphists.

Our methods are designed to support specific expert needs and include visual retrieval and classification of glyphs. This is a collaboration with Maya archaeologists from the University of Bonn and computer scientists from the University of Geneva, and is supported by the SNSF-DFG project MAAYA³⁸ project (Multimedia Analysis and Access for Documentation and Decipherment of Maya Epigraphy). This work has received recent press coverage^{39 40}.

xii

³⁸ www.idiap.ch/project/maaya/

⁴⁰ www.nzz.ch/wissenschaft/archaeologie/google-translate-fuer-maya-hieroglyphen-1.18639959

³⁹ www.letemps.ch/sciences/2015/11/02/un-google-translate-ecriture-maya

4. Perception and Activity Understanding

OVERVIEW

Head: Dr. Jean-Marc Odobez (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 (mainly vision and audio). This entails the investigation of fundamental tasks like the representation, detection, segmentation and tracking of people, the characterization of their state, and the modeling of sequential data and their interpretation in forms of gestures, activities, behavior or social relationships. These tasks are addressed through the design of principled algorithms extending models from computer vision, statistical learning, or multimodal signal processing. In particular, the group has focused on the recent years on the exploitation of probabilistic graphical models as one of its main modeling tools. Another objective of the group is the investigation, whenever possible, of unsupervised adaptation methods to leverage either the availability of potentially large amounts of test data from which exploitable statistics can be extracted, or of prior models that can help the soft tagging of test data, or of both. Surveillance, traffic analysis, analysis of behavior, human-robot interfaces, and multimedia content analysis are the main application domains.

In 2015, the group was composed of 2 post-doctoral members, 3 PhD students, and 1 master student.

KEY SCIENTIFIC OUTPUTS

The group is well known for its work on video sequence analysis, probabilistic tracking and non-verbal behavior extraction. In 2006, it developed a unique video stabilization algorithm

that is currently applied automatically on all super-8 and videos digitized by the Cinetis company. In 2007, the Klewel company was created using its slide OCR technology. More recently, the work on 3D face and gaze tracking is leading to the creation of the Eyeware start-up. During the period 2011–2015, it published 20 journal papers and around 50 conference papers, and received 1 best paper and 1 best student paper awards. The group has pioneered work on unsupervised surveillance scene activity, visual attention modeling, and its recent work on gaze estimation from cheap RGB-Depth sensors is under patenting. With their LIUM partner in the EUMSSI project, the team ranked first out of 8 teams at the MediaEval Person discovery challenge in 2015.

MULTIPLE OBJECT TRACKING

The recent work in this research line has resulted in an enhanced tracking-by-detection system with several important characteristics that make it successful in challenging conditions, and which are illustrated in Fig. 11. Given the individual object detections (humans, faces) at each instant, the problem is set as a graph partitioning problem (within a Conditional Random Field CRF framework) such that detections corresponding to the same object are associated together based on fine grained multicue measurements (position, image-based motion, multiple patch colors, set of SIFT descriptors). Important novel features of our model are the following. Links between detections can be a few seconds apart, rather than only a few frames apart. All links are used in defining the cost function, not only those between consecutive nodes from the same track. This allows assessing a solution using both short and long-term links. Importantly, pairwise cue measurement costs (defined as similarity/dissimilarity likelihood ratio) are sensitive to the time interval between two detections, integrate multi-cue association reliability factors, and all likelihood distribution parameters are learned in an unsupervised fashion, thus accounting for the data, scene, and cue specificities while avoiding tedious annotations. Results equal or better than the state-of-the-art were obtained on several benchmark data of the field, regarding both human and face tracking. The work received the best paper award at the annual IEEE Performance Evaluation of Tracking and Surveillance (PETS) workshop. The software has been licensed for research purposes to a large leading IT company active in camera manufacturing.

XIII

²⁶ <http://www.idiap.ch/project/sensecityvity>

²⁷ <http://www.conacytprensa.mx/index.php/centros-conacyt/2143-sensecityvity-estudio-de-la-geografia>

²⁸ <http://conacytprensa.mx/index.php/>

centros-conacyt/5198-crece-proyecto-sensecityvity-en-mexico-nota

²⁹ <https://www.youth-night.ch/>

³⁰ <http://www.idiap.ch/project/mcsc>

³¹ <https://www.youtube.com/>

watch?v=UZT4bfVnL9Q

³² <http://cana19.ch/comportements-humains-lidiap-sonde-les-smartphones/>

³³ <http://actu.epfl.ch/news/an-app-to-steer-clear-of-kenya-s-road-hazards/>

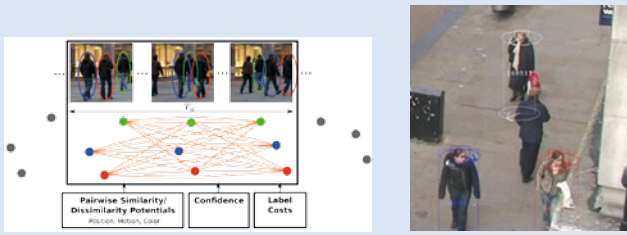


Figure 11
 Left: the multi-person tracking framework. Human detections at each frame are associated based on long-term multi-cue and time-interval sensitive discriminative costs. Right: estimated head and body orientation of people.

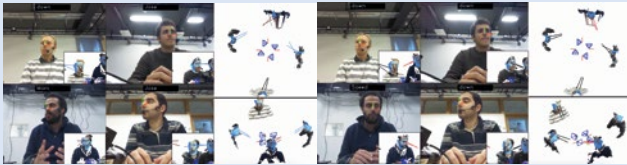


Figure 12
 Heads are tracked, eye localized and aligned before inferring gaze and attention towards people or other semantic labels. The method analyses all information in 3D (cf different renderings), allowing an easy inference of attention cues.

BEHAVIOR CUE EXTRACTION IN SURVEILLANCE SCENES

We have worked in recent years on body and head pose estimation (see Fig. 11, Right) which are important cues for activity analysis. Our work addressed the problem through the joint adaptation to new scene of head and body pose classifiers. The method accounted for the different coupling (between head and body pose due to anatomical constraints, between motion and body pose conditioned on a walking status). It also handled the discrepancy problems resulting from the mismatch between training and test data, using a synchronized manifold alignment approach which learn two feature projection mappings enforcing train and test features in the projected space to form a single smooth pose manifold. The method thus leverages at several modeling places the availability of weak pose labels (e.g. from people motion) in the test data to learn better pose classifiers.

GAZE AND ATTENTION MODELING

Our long-standing work on gaze and attention been continued in the context of the SNSF funded project G3E, which more specifically addresses gaze extraction from cheap RGB-D (depth) cameras like Kinect. Building on our the

head-pose free gaze estimation framework, we have proposed several extensions to enhance the performance at different levels (gaze estimation, software system, usability) and gain robustness against low eye image resolution and imprecise pose estimation, e.g. through synchronized delaunay triangulation eye alignment. The figure 12 illustrates the result of applying our tools for gaze coding in multiparty situations.

NON-VERBAL BEHAVIOR EXTRACTION AND HUMAN INTERACTION ANALYSIS

In addition to gaze, our work on non-verbal behavior cue extraction and human interaction modelling has continued in the context of the UBIMPRESSED project. In particular, we have worked on the 23 / 69 2 RESEARCH GROUPS 2015 Idiap Annual Report – March 2016 Figure 13: 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. recognition of head gestures like nods which by conveying agreement, emphasis, or doubts, constitute a potential strong cue for the inference of first impressions in natural interactions. To handle the recognition of subtle nods and be immune to the camera view point, we have proposed a novel gesture detection method that leverages on our 3-D face tracker to compute (i) view-invariant relative head rotation from which rich dynamics oscillatory features based on Fourier transform are derived, (ii) rotation axis features allowing to distinguish oscillations coming from the trunk or from head oscillations around the neck. The method has been extended to the multi-class head gesture case.

In collaboration with the 'Speech and Hearing group' of KTH, we have worked on the analysis of group interaction and more specifically studied the classification of silent participants in three categories (attentive, side participants, bystander). Turn-taking, gaze tracking and head nodding cues have been automatically extracted using our tools and used to better characterize the engagement of participants. The study showed that many of the audio-visual cues (mutual gaze, backchannels) distinguishing listeners in dyadic interactions were also holding in multi-party interactions.

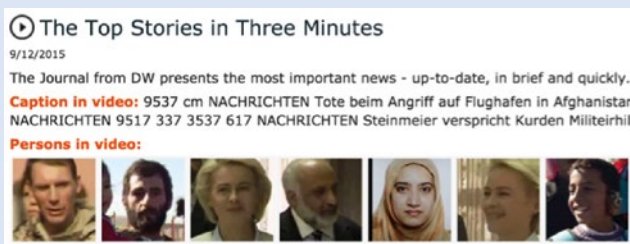


Figure 13

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 allows to monitor who said what and when and is thus important for the design of search and fast browsing tools of personal data collection or broadcast programs, as illustrated in Fig. 13. This is one of the challenge that we are addressing in the EU funded EUMSSI project, whose global aim is to provide journalists with text and audio visual search tools for event understanding. This year, we have improved the robustness and efficiency of the overall processing pipeline, with an emphasis on the face diarization part (face detection, tracking and clustering). In addition, we have evaluated the benefit of state-of-the-art face representations (i-vectors, neural-network), and shown that in the diarization context, matching representations based on SURF descriptors are still necessary to handle the large variety of pose and expression found in the media documents.

VISUAL ANALYSIS OF MAYA HIEROGLYPHS

In collaboration with the Social computing group, we are working on the automatic visual analysis of hieroglyphs from the ancient Maya codices, see Fig. 10.

5. Robot Learning & Interaction

OVERVIEW

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

GROUP OVERVIEW

The Robot Learning & Interaction group, created in May 2014, focuses on human-centric robot applications. The group targets the problem of transferring skills to robots with intuitive human-robot teaching interaction. The goal is to endow robots with learning systems that can acquire rich motor skills adaptively without requiring the users to program the robots through a computer language. The objective is to develop probabilistic models of movements and behaviors for robots evolving in unconstrained environments and sharing the same workspace as humans. In these applications, the models serve several purposes (recognition, prediction and synthesis), and are shared by different learning strategies (imitation, emulation, exploration and self-refinement). The Robot Learning & Interaction group in 2015 was composed of one postdoctoral fellow, one PhD student, one visiting PhD student and two MSc students.

TASK-ADAPTIVE MOVEMENT LEARNING

The group investigates the problem of analyzing and generating movements by relying on demonstrations collected in diverse situations, such as recordings with different positions of objects in a manipulation task.

KEY SCIENTIFIC OUTPUTS

We developed an approach modeling the observed motion in multiple coordinate systems (frames of reference). The different projections of the collected data reveal regularities that are exploited to generalize and adapt the task to new situations in a probabilistic manner. Notably, the retrieved variability and correlation information in the different coordinate systems is used to regulate the stiffness of the robot in order to generate safe and natural movements that concur with the task objectives. Fig. 14 illustrates the proposed approach. It enables new haptic communication capability that has great potential in human-robot collaboration.

Aspiring to reproducible research, the group released open source codes of the proposed approach and published an accompanying tutorial article. The source codes are available in two versions (either 25 / 69 2 RESEARCH GROUPS 2015 Idiap Annual Report – March 2016 C++ or Matlab/GNU Octave)⁴¹.

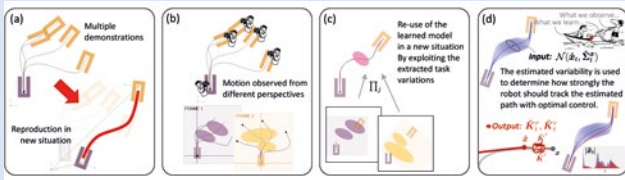


Figure 14

Task-parameterized probabilistic model to synthesize movements in new situations. (a) Observation of the task in different situations and generalization to new contexts. Multiple demonstrations help at extracting the structure of the task. (b) Probabilistic encoding of continuous movements in multiple coordinate systems. (c) Exploitation of variability and correlation information to adapt the motion to new situations. With cross-situational observations of the same task, the robot can generalize the skill to new situations. (d) Computation of the underlying optimal control strategy to reproduce the observed behavior.

LEARNING FROM DEMONSTRATION FOR SEMI-AUTONOMOUS TELEOPERATION

Robots are getting closer to humans in diverse manners, sharing the users' space (humanoids at home and in manufactures, robots as co-workers), being in direct contact with the users (exoskeletons, prosthetics, rehabilitation) or inside their bodies (surgical robots). Such human-centric robot applications require a tight integration of learning and control. This connexion can be facilitated by a shared representation of the tasks and objectives to achieve in a probabilistic form. In human-robot collaboration, such representation can take various forms. Movements must be enriched with perception, force and impedance information to anticipate the users' behaviors and generate safe and natural gestures.

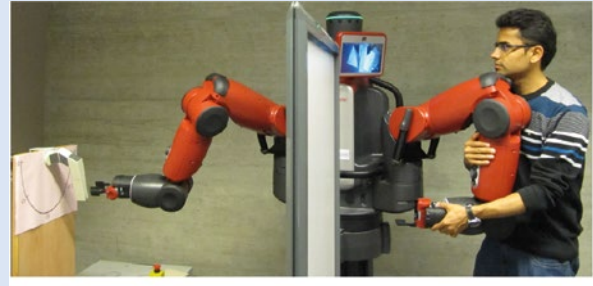


Figure 15

Teleoperation with the Baxter robot, in which one of the two arms is used as a haptic input device, and the other arm is used to imitate the movement in a semi-autonomous manner.

KEY SCIENTIFIC OUTPUTS

The group proposed an approach to extend learning from demonstration techniques to assistive tasks and semi-autonomous teleoperation. The approach is exploited within the DexROV European Project, in which a bimanual underwater robot is distantly controlled by a user wearing an exoskeleton with force feedback. The transmission delays are handled by treating teleoperation as a joint classification, prediction and synthesis problem, all relying on a shared model encoding a set of motion, coordination and impedance primitives that can be assembled in sequence and in parallel. On the teleoperator side, the model is used to recognize existing movement/behavior primitives that were used in previous missions and predict their most likely outcomes. On the manipulator side, the same model is used to generate autonomously a movement/behavior until a new signal from the teleoperator is available. Fig. 15 shows the semi-autonomous teleoperation setup developed at Idiap to study such challenge.

⁴¹ www.idiap.ch/software/pdlib/

6. Artificial Cognitive Systems

OVERVIEW

Head: Prof. Barbara Caputo
(PhD, Royal Institute of Technology, Sweden, 2005)

GROUP OVERVIEW

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

The Artificial Cognitive Systems group in 2015 was composed of 1 head of group and 2 PhD students.

KEY SCIENTIFIC OUTPUTS

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

of the value of our benchmarking activities, during 2013 and 2014 Prof. Caputo was the main organizer of ImageCLEF. In 2014, Prof. Caputo received an ERC Starting Grant.

SEMANTIC SPATIAL UNDERSTANDING OF INDOOR PLACES

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

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

Over the last year, we casted the semantic spatial modeling problem into that of learning from multiple cues. Our contributions have been a principled online Multi Kernel Learning algorithm able to combine optimally multiple features while providing theoretical guarantees on the expected performance, and a general learning to learn algorithm, based on the cue integration paradigm, able to leverage over prior

⁴² <http://www.idiap.ch/project/ninapro/>

knowledge in an extremely flexible and adaptive manner, over a wide range of application scenarios. This last algorithm, the first of its kind in the literature, achieves state of the art results on benchmark databases widely used in the literature and is currently implemented in a running robot system.

ADVANCED HAND PROSTHESES

The work on advanced hand prostheses has been initiated in 2011 with the SNSF funded NINAPRO⁴² project. Its goal is to develop a family of algorithms able to significantly augment the dexterity, and reduce the training time, for sEMG controlled prosthesis. Indeed, the state of the art in hand prosthetics does not offer more than 2–3 degrees of freedom and a very coarse control of the force, as there is no haptic feedback. Patients interface with the prosthesis via surface electromyography (sEMG), recorded using surface electrodes. Learning how to control the device through many input sEMG channels is a long and difficult process for most patients, that therefore settles for limited and very simplified movements (open/close).

This contrasts with recent advances in mechatronics, thanks to which mechanical hands gifted with many degrees-of-freedom and force control are being built. There is a need for prosthetic hands able to naturally reproduce a wide amount of movements and forces, while at the same time requiring a lower effort in learning how to control hand postures. This goes beyond mechatronic dexterity: the real challenge is how to provide patients with a cheap, easy and natural way of controlling the prosthesis. By testing its findings on a very large collection of data, NINAPRO has the ambition to pave the way for a new generation of prosthetic hands. The work is organized along four themes: data acquisition and analysis (Theme 1), augmented dexterity: posture classification (theme 2), augmented dexterity: natural control (theme 3) and adaptive learning (Theme 4). Prof. B. Caputo is the coordinator of the project, and the principal investigator in the Themes 2 and 4 (Figure 16).

The objective of theme 2 (augmented dexterity) is to push the current state of the art in prosthetic hand posture classification from handling a maximum of 12 postures up to 40–50. Over the last year, we have demonstrated the importance of using multi modal features to increase the robustness of posture

classification and therefore of actual control of the prosthesis, by merging information acquired from sensors applied on the stump of the user with those recorded by an accelerometer mounted on the prosthesis. This work has resulted in a publication on the international journal IEEE Transaction on Neural Systems and Rehabilitation Engineering.

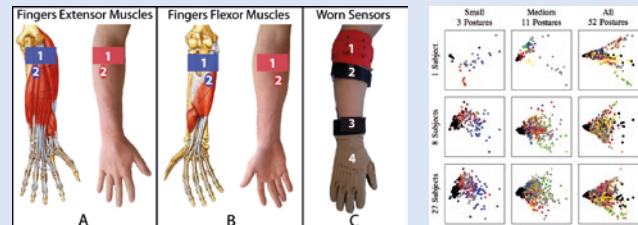


Figure 16

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

The goal of theme 4 (adaptive learning) is to develop learning algorithms to better interpret the sEMG signals acquired from the patients, with the ultimate goal of boosting the learning process necessary for the patient to effectively use the prosthesis. Our plan is to build pre-trained models of various data postures, on the data acquired in theme 1, and adapt these general models to the needs of individual users as new data will become available using adaptive online learning methods. Over the last year, we did build over our results obtained in the large-margin classification framework, and we explored how to maintain the stability and effectiveness of our method when in presence of a very large number of prior models. This would be indeed the case when leveraging over priors derived from the actual experience of intact subjects, from which it is relatively easy to acquire large amounts of data and hence build strong prior models. We developed a new version of our previous algorithm able to scale well over a very large number of priors in terms of computational efficiency, while preserving a very high accuracy in terms of hand posture recognition.

⁴² www.idiap.ch/project/ninapro/

7. 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 system modelling. 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 high-fidelity experiments and/or numerical simulations depending on a number of inputs including both controlled design parameters and uncontrolled environmental variables.

Uncertainty Quantification comprehends a set of statistical approaches with the aim to model, elicit and characterize uncertainties arising in the study of complex systems, be it relying on data, numerical simulation and/or on expert knowledge. Optimal Design essentially deals with the choice of controlled design parameters relying on global optimization methods. In particular, elaborated model evaluation strategies are needed for making optimal decisions, both for deterministic numerical simulations and for design under uncertainty.

More specifically, the main focus of the UQOD group is on stochastic models and adaptive design of experiments for optimization, inversion, and related problems. Typical stochastic models studied within UQOD encompass spatial random fields and times series, notably with Gaussian Processes but also with an ongoing interest for non-stationary extreme value modelling.

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. In its initial form it was composed of one head of group, one IMSV postdoc and one IMSV PhD student.

KEY SCIENTIFIC OUTPUTS

Current contributions include efficient algorithms for Bayesian set estimation, notably for estimating and quantifying uncertainties on overcritical parameter regions and other implicitly defined sets of interest with Gaussian random field models. Other recent results deal with the interplay between the choice of the covariance kernel and degeneracy/invariances properties of Gaussian random field 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 simulation-based optimal placement of decontamination pumping wells.

DETERMINISTIC AND STOCHASTIC FUNCTION MODELLING WITH KERNEL METHODS

Kernel methods, and in particular Gaussian Process (GP) / Gaussian random field models, have become quite popular in the last decades for approximating costly-to-evaluate functions relying on evaluations at a finite set of points. Often, kernels used in such methods are taken among off-the-shelf families, typically isotropic or anisotropic stationary kernels, e.g. from the Matérn family. On the other hand, in a number of problems some expert knowledge is available about the function to be approximated, e.g. being zero-mean, possessing given symmetries and invariances, being harmonic, etc. It turns out that these properties can be cast as degeneracies under linear operators, and so kernels can be designed that incorporate such structural properties, leading both to a predictor honouring the expert knowledge but also to a posterior distribution of elements honouring it as well. Previously and recently obtained results along this line are currently being resumed, with the aim to explore ways of putting more expressiveness in GP models. In particular, links to global sensitivity analysis are considered, paving the way to original high-dimensional GP-based function emulation approaches.

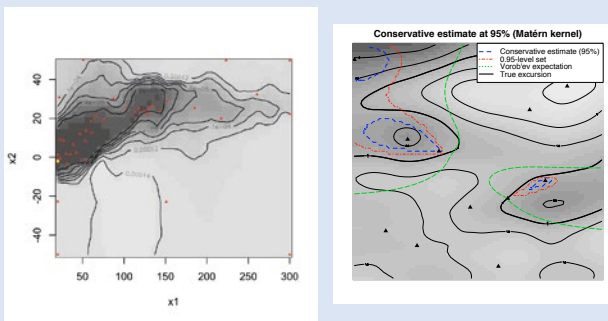


Figure 20
 Surrogate-based minimization of the highest concentration (over time) escaping the right outlet of a domain as a function of coordinates of a remediation pumping well (left). Right: conservative excursion set estimation under a Gaussian random field prior.

Also, as illustrated in Fig. 20 (left panel), the group's know-how in surrogate-based strategies for global optimization is put to good use with a simulation-based application in well placement for decontamination set up together with hydro-geologists from the Universities of Lausanne and Neuchâtel. Here a response quantifying how bad some contaminant escapes from a given area is minimized as a function of the location of a candidate pumping well. The yellow point stands for the optimal placement (obtained from a very time-consuming extensive grid search) while the red points are points explored by a state-of-the-art global optimization algorithm relying on a GP model. Comparisons with other algorithms and extensions to more input variables (e.g., placing several wells) are currently under study. Also, extensions of GP-based optimization algorithms to the case of larger number of points than what basic GP models can accommodate is part of current research topics of the group.

EFFICIENT ALGORITHMS FOR PROBABILITY OF EXCURSION AND CONSERVATIVE SET ESTIMATION

In a number of practical problems from natural sciences or the industry that involve costly-to-evaluate functions, the goal is not exactly to optimize the function but more precisely to locate regions of the input space leading to a response above or below a given threshold, that correspond for instance to a dangerous or abnormal behaviour of the considered system. Gaussian random field models have been increasingly used to approximate such excursion sets, and also to quantify and reduce uncertainties on them by notably using stochastic simulation techniques and sequential evaluation strategies. Recently, approaches have been proposed for conservative estimation of such sets; conservative here means that under the considered model the actual excursion set is with a given probability

completely inside (or outside, depending on the way the problem is formulated) the estimate returned by those methods. Hence deriving such estimates necessitates to calculate the probability that the maximum of a high-dimensional vector (say, in a discretized case) remains below or above a given threshold. While algorithms have been proposed to calculate such probabilities, we have been working on a special Monte Carlo technique to reduce the overall computational cost of this kind of calculation, leading in particular to speed ups in the computation of conservative excursion set estimates. The right panel of Fig. 20 presents the 95% conservative excursion set estimate obtained for a Gaussian random field simulation test case, and a comparison with the actual excursion set and other candidate estimates which level of confidence is substantially lower than 95%.

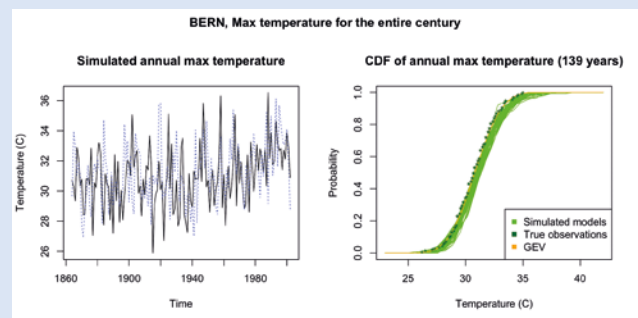


Figure 21
 Left: Two realizations of annual max temperature at Bern for the horizon 1864-2002 under a proposed stochastic model. Right: Corresponding simulated (longitudinal) empirical CDFs together with the empirical CDF of observed yearly maxima and a Generalized Extreme Value distribution fitted to them.

xx

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

Extreme value analysis is a key statistical approach to a number of phenomena from nature, economic and industry activities. In order to assess risks and make design or policy decisions accordingly, it is of importance to understand and forecast how extreme some 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 numerous decades, and statistical tools are available to fit so-called "Generalized Extreme Value" (GEV) distributions to observed maxima and quantify how severe and with what probability extreme events may occur. However, the most spread usage of GEV assumes a sample of independent identically distributed maxima in order for the underlying mathematical theory to be valid. Some attempts have been done to fit GEV distributions in non-stationary conditions, but a number of practical and research questions remain open. This poses serious challenges to climate scientists and beyond, as phenomena turn out to exhibit dependences and also to change over time. How then to speak about extremes, e.g. about return levels, based on non-stationary sequences of observations?

In approaches currently being developed within the UQOD group and in collaboration with the Oeschger Center for Climate Change Research we notably consider times series data at several scales, departing from the plain block-wise maxima approach where distributions are fitted to a sample of yearly maxima while usually neglecting what is happening at a finer temporal frequency. Our investigated algorithms for characterizing and forecasting extremes of times series essentially rely on state-of-the-art time series decompositions, on Monte Carlo simulations, and on statistical learning prediction algorithms.

8. 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 multi-modality images.

In 2015, the Computational Bioimaging Group was composed of the head of group and one trainee.

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.

QUANTIFYING 3D BLOOD-FLOW AT THE MICROSCOPIC LEVEL

Conventional fluid flow estimation methods for in vivo optical microscopy are limited to two-dimensions and are only able to estimate the components of flow parallel to the imaging plane. This limits the study of flow in more intricate biological structures, such as the embryonic zebrafish heart, where flow is three-dimensional. To measure three-dimensional blood flow, we have proposed, together with UCSB student and Idiap intern Kevin Chan, an algorithm to reconstruct a 3D, divergence-free flow map from multiple 2D flow estimates computed from image stacks captured from different views. This allows us to estimate the out-of-plane velocity component that is normally lost with single-view imaging. This paper describes our 3D flow reconstruction algorithm, evaluates its

performance on a simulated velocity field, and demonstrates its application to in vivo cardiac imaging within a live zebrafish larva.

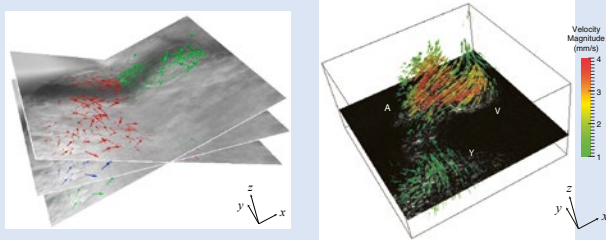


Figure 22

Left: We acquired focal stacks of heart in a zebrafish larva from three views: -18° , 0° , 18° relative to the xy -plane and rotated about the y -axis. At each view, 2D optical flow was used to estimate velocity vectors in each plane. For visibility, a single slice is shown for each view. Right: We combine 2D flow estimates from three different views to recover a divergence-free 3D velocity map of blood flow through the heart (A: atrium, V: ventricle, Y: yolk). Adapted from Chan et al., ISBI 2015.

IMAGE PROCESSING METHODS FOR MULTI-VIEW MICROSCOPY

Optical microscopy offers the unique possibility of imaging living samples in 3D under conditions similar to their native state. However, most techniques suffer from anisotropic 3D resolution (worse in the axial than the lateral dimensions). This makes quantification of images more difficult. Fusion of multiple datasets acquired from different view points has the potential to overcome this limitation. Together with UCSB graduate student Nikhil Chacko we developed an algorithm to spatially register two volumetric datasets related via a rigid-body transform and degraded by an anisotropic point-spread-function (PSF), as a first step towards solving the fusion task. Our automatic registration approach, a multi-scale pixel-based registration method accounts for the anisotropic image formation process by including a re-blurring operation before minimizing the mean squared intensity difference between a reference and a test volume.

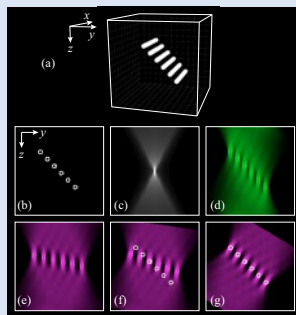


Figure 23

(a) 3D perspective of the object used in simulations. (b-g) Maximum intensity projection (MIP) in the yz plane of f , h (after gammacorrection of $= 0:3$), fR , fT , the registered results using the traditional, and our proposed approach, respectively. For comparison, (b) has been overlaid on the registered results shown in (f) and (g). Scale bar is 25m. Adapted from Chacko et al. ISBI 2015.

9. Biometric Person Recognition

OVERVIEW

Head: Sébastien Marcel (PhD, University of Rennes, France, 2000)

GROUP OVERVIEW

The Biometrics group investigates and develops novel image-processing and pattern recognition algorithms for face recognition (2D, 3D, and near-infrared), speaker recognition, vein recognition, anti-spoofing (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. The Biometric Person Recognition group in 2015 was composed of 1 head of group, 2 scientific collaborator, 4 postdocs, 2 PhD students, few interns and visitors.

KEY SCIENTIFIC OUTPUTS

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

FACE AND SPEAKER RECOGNITION

We leveraged prior work on distribution modeling for part-based face recognition using session variability modeling techniques. Session variability modeling aims to explicitly

model and suppress detrimental within-class (inter-session) variation. We examined two techniques to do this, inter-session variability modeling (ISV) and joint factor analysis (JFA), which were initially developed for speaker recognition. Next, we explored Total Variability modeling (TV), so called i-Vectors originally proposed for speaker recognition, for the task of face recognition, and we reached boosted performance for the task of bimodal face and speaker recognition (Figure 24).

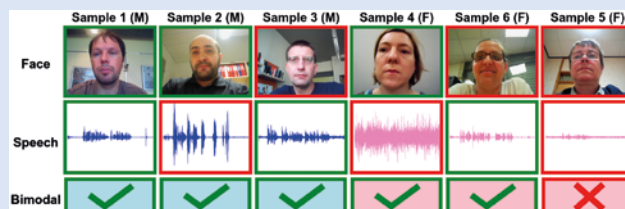


Figure 24

An illustration of face and speaker recognition.

We also developed recently a scalable formulation of Probabilistic Linear Discriminant Analysis (PLDA). PLDA is a probabilistic model that has been shown to provide state-of-the-art performance for both face and speaker recognition. However, it has one major drawback, at training time estimating the latent variables requires the inversion and storage of a matrix whose size grows quadratically with the number of samples for the identity (class). We developed a scalable derivation which is theoretically equivalent to the non-scalable solution and so obviates the need for a variational approximation. The main advantage of the proposed approach is a gain in performance when using more training samples per identity (class), which is made possible by the proposed scalable formulation of PLDA. As a consequence, on July 31 2014 the **biometrics group was ranked 1st** among 130+ participants to the NIST i-Vector Machine Learning challenge on Speaker Recognition (ivectorchallenge.nist.gov) using this novel PLDA combined with an efficient hierarchical clustering.

ANTI-SPOOFING

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

the existence of very sophisticated biometric systems nowadays, the task of implementing anti-spoofing schemes for them has attracted much less attention. Recently, we advanced the state-of-the-art in anti-spoofing to printed photo attacks in face recognition by proposing public benchmarks and effective solutions based on texture modeling, motion detection or spatio-temporal pattern recognition. We also organized the first International Competition on the topic.

REPRODUCIBLE RESEARCH

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

REPRODUCIBLE RESEARCH

Within the EU BEAT project⁵⁰, coordinated by Idiap, we built a biometry-independent platform for Biometrics research, development and certification. By making use of such a system, academic or industrial partners enable users to easily compare results from distinct algorithms or/and parameterizations with minimal interaction using one or potentially many biometric traits. Here are three practical examples:

1. A teacher asks students to implement a specific illumination normalization algorithm in a preconfigured face recognition tool-chain. The teacher would like to receive the examination results from his group of students;
2. A researcher (PhD student or Professor) wishes to test new

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

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

One can use the platform to generate attestation (a.k.a. certificates) to any produced experimental results. Attestation can be used to support formal certification processes or improve evidence on published academic articles (Figure 25).

XXIV

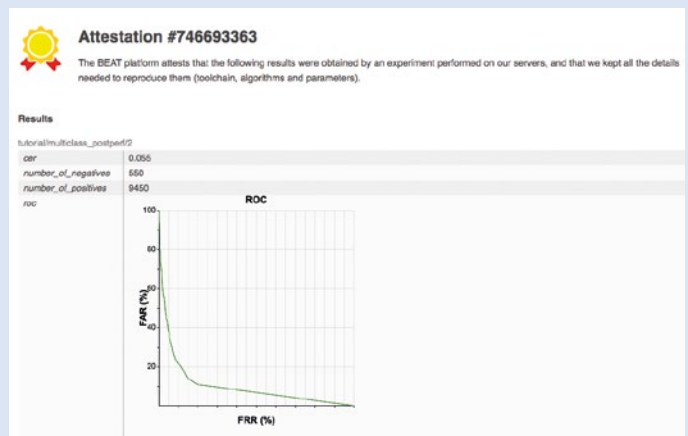


Figure 25
A snapshot of a BEAT attestation.

SWISS CENTER FOR BIOMETRICS RESEARCH AND TESTING

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

⁵⁰ www.beat-eu.org/

projects and to push for industry-driven research and testing in biometrics. The center will mainly promote two instruments: (1) the coordination of a cooperative research consortium, and (2) the deployment of the BEAT evaluation platform. These instruments were presented on November 2014 during a kick-off to researchers and stake-holders in biometrics (40+ participants). Consequently to this kick-off, we decided to elevate the visibility of the cooperative research consortium to a more European level. We concluded an agreement with the European Association for Biometrics (composed of 160 members) to transform the cooperative research consortium initiated in Switzerland into an European cooperative research consortium but still coordinated in Switzerland by the center. We expect to attract a critical mass of partners to fund the European cooperative research consortium. Regarding the BEAT evaluation platform, we were able to speed-up its development for a deployment late 2015. As a consequence, we have been contacted already by institutions and companies to use the platform for evaluation and hence we plan to monetize the platform by providing commercial licenses. Finally, all these activities related to the center attracted the attention of large companies and led to a license agreement with Safran Morpho, a research agreement with Google and an US IARPA project.

10. Natural Language Processing

OVERVIEW

Heads: 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 important 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 semantic relatedness, disambiguation of pronouns or connectives, keyword extraction, or sentiment analysis – and guides these improvements so that they become useful to the two applications, MT and IR. Regarding the first domain, the NLP group combines text-level processing techniques with phrase-based statistical MT systems to improve translation quality. Regarding the second task, the NLP group exploits content and sentiment analysis to improve document and multimedia recommendation, in realtime or not, over networked repositories.

xxv

The NLP group had in 2015 the following members: the head of the group; two postdoctoral students; four PhD students (including one visitor); and one intern. A PhD thesis was defended in autumn 2015 and another one is scheduled for spring 2016.

KEY SCIENTIFIC OUTPUTS

We have demonstrated that semantic and coherence-based analyses of discourse connectives, verb tenses, noun phrases and pronouns in multi-sentence texts are beneficial to statistical MT, an innovation that has been met with interest from the community. We have integrated novel algorithms for keyword extraction and document ranking, which reward the diversity of results, into a system that recommends Wikipedia pages in real-time to conversation participants. We have shown that sentiment analysis of user comments improves recommendation, and that it is possible to predict individual sentiment aspects and relate them to text excerpts that explain the predictions using multiple instance regression; moreover, the learned

weights of the excerpts are predictive for segmentation and summarization of comments. These achievements have been demonstrated at competitive evaluations such as the ACM Multimedia Grand Challenge, the MediaEval hyperlinking task, and the EMNLP DiscoMT shared task on pronoun translation. They have also been used for technology transfer with startup companies.

TEXT-LEVEL MACHINE TRANSLATION

We focus on a problem that is not often targeted in the current statistical machine translation (SMT) paradigm: the translation of words that depend on long-range relationships between sentences. The main examples are discourse connectives (which have been our main focus in 2010–2014), pronouns and noun phrases (our more recent focus), and verb tenses (studied in 2013–2014). Connectives and verb tenses have been mainly addressed in the COMTIS SNSF Sinergia project⁵¹, and completed in the first year of the MODERN SNSF Sinergia project⁵². The main focus of MODERN is on the coherent translation of noun phrases and pronouns. Both projects are coordinated by the head of the NLP group, with teams from Geneva, Zürich and Utrecht.

xxvi

The negative impact of wrong pronoun English/French translation is illustrated with the following example, translated by online SMT. In the source, the pronoun *it*₃ refers to the mouse, which is feminine in French, but is translated by the SMT into masculine *le*₃, making the readers think that the cat (masculine in French) was hurt and that the writer intends to kill it.

Source text:

My cat brought home a mouse that he hunted, and *it*₁ was not dead but *it*₂ was mortally wounded. What is the best way to kill *it*₃ humanely?

Machine translation:

Mon chat a ramené à la maison une souris qui il a chassé, et **il*₁ était pas mort, mais **il*₂ a été mortellement blessé. Quelle est la meilleure façon de **le*₃ tuer humanement?

Correct:

Mon chat a ramené à la maison une souris qu'il a chassée, et *elle*₁ n'était pas morte, mais *elle*₂ était mortellement blessée. Quelle est la meilleure façon de *la*₃ tuer humanement?

In 2015, we addressed the problem of noun and pronoun coherence in translation as follows. (1) We enforced the consistency of translations of nouns that appear first in a compound and then alone, and demonstrated that this simple strategy improves their translation from German and Chinese into English. Later, we have generalized this approach with a classifier that hypothesizes, when the same noun is repeated in a source text, whether the first translation should replace the second one, or viceversa, or no such post-editing should be done. (2) We designed two systems for the DiscoMT 2015 shared task⁵³. The system for pronoun prediction does not rely on coreference resolution, and uses instead a large set of source-side and target-side features to train a pronoun predictor. The system for pronoun-focused translation combines two knowledge sources: grammatical constraints from the hypothesized coreference links, and candidate translations from an SMT decoder. The latter system was ranked first among seven participants at the DiscoMT shared task. (3) Improving on this successful system, we designed a solution for coreference-aware decoding, which uses a coreference model with information about the anaphoric pronouns and their likely antecedents. Scores were thus improved on the DiscoMT 2015 task. (4) To evaluate automatically the translation of pronouns, we compared them automatically with those in a reference translation with the APT metric (Accuracy of Pronoun Translation), and showed that it correlates well with human judgments.

Moreover, to explore text-level SMT coherence even further, we conducted experiments that re-rank the hypotheses of an SMT system based on their coherence with the topics inferred from previous sentences, as modeled using a topic-aware keyword extraction method previously designed in the NLP group. Similarly, re-using previous work on sentiment analysis, we explored sentence-level and word-level constraints on the sentiment of target words to re-rank MT hypotheses. Finally, through the work of a visiting PhD student, we conducted experiments with speech-based editing of SMT output, showing that features such as ASR confidence and word length are relevant to decide which hypothesis (ASR or MT) to select.

⁵¹ 2010–2013, see www.idiap.ch/project/comtis

⁵² 2013–2016, see www.idiap.ch/project/modern

⁵³ The 2nd Workshop on Discourse and Machine Translation (at EMNLP), co-chaired by the head of the NLP group.

TEXT ANALYSIS FOR MULTIMEDIA RECOMMENDATION

We developed methods for multimedia information retrieval and recommendation based on the analysis of content and on the networked structure of social media repositories. This work is supported by the InEvent EU project⁵⁴, TheArk foundation⁵⁵, and the CTI/KTI⁵⁶. A PhD thesis started within the IM2 NCCR and the AMI/AMIDA EU projects⁵⁷ was defended in 2015, on diverse keyword extraction from conversations, and the use of these keywords to perform just-in-time document recommendation in conversations. The theoretical contributions have been integrated into an end-to-end system (ACLD, the Automatic Content Linking Device), for which a brainstorming task and evaluation metrics have been designed; the system was favorably rated for usability in evaluation experiments.

At the end of 2015, a second PhD thesis had a successful private defense at EPFL, mainly supported by the inEvent project. The thesis proposes novel methods based on machine learning to model user sentiment and preferences for information filtering systems. In 2015, the problem of modeling user attitude toward aspects of items was solved using a weighted multiple-instance learning model. The saliency weights learned by the model are useful features for review segmentation and summarization. Experiments were conducted on real-world data collected from online social networks or Web communities such as TED, Flickr, Vimeo, or Audible.

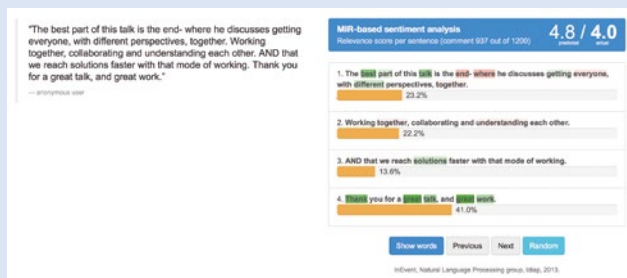


Figure 26
Sentence relevance weights hypothesized using multiple-instance learning for predicting the sentiment of a user comment related to a TED talk.

Working on personalized search and recommendation for information watch with Faveeo SA, we designed implicit query expansion solutions to increase the relevance of the ranked list of results for complex queries. Within a hybrid semantic search framework, we proposed a two-step retrieval and reranking strategy and compared two methods for query expansion, one based on word co-occurrence in the document repository, and the other one using semantically-related expansion terms drawn from external resources. We evaluated our methods and compared them with the Faveeo baseline by crowdsourcing user-centric tasks on three evaluation dimensions: relevance, diversity, and utility, as illustrated in Figure 27. The results show that our proposed hybrid semantic search approach achieves better results in terms of relevance and overall utility than the initial baseline, with word co-occurrence outperforming semantic query expansion. In addition, we extended word co-occurrence methods to the recommendation of explicit keywords for query expansion to users, and also applied implicit query expansion to spoken queries in the ACLD framework.



Figure 27
Sentence relevance weights hypothesized using multiple-instance learning for predicting the sentiment of a user comment related to a TED talk.

⁵⁴ www.inevent-project.eu

⁵⁵ Technology transfer project with Unono.net, a university social network.

⁵⁶ Technology transfer project with Faveeo SA, an information watch company.

⁵⁷ See www.im2.ch, www.amiproject.org and www.amidaproject.org, finished in 2013.

Projects in progress during 2015

PROJECTS FUNDED BY EUROPEAN AGENCIES

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

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

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

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

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

Name SIIP (Speaker Identification Integrated Project)
Funding FP7 – SRTEP – SEC
Coordinator Verint System Ltd
Duration 01.05.2014 – 30.04.2018
Partner(s) Cassidian SAS, Laboratorio Di Scienze Della Cittadinanza, Synthema S.R.L, Idiap Research Institute, Police Service of Northern Ireland, Loquendo SpA, Singarlogic Anonymos Etairia Pliroforiakon Systimaton & Efarmogon Pliroforikis, International Biometric Group (UK) Limited, Ministério da Justiça, Lisboa, The Internation Criminal Police Organization, Sail Labs Technology AG, Green Fusion Limited (Data Fusion Internation), ok-2go, University of Warwick, Inov Inesc Inocacao, Instituto De Novas Tecnologias, Rijksuniversiteit Groningen

PROJECTS FUNDED BY SWISS AGENCIES

Name **MCM-FF** (Multimodal Computational Modeling of Nonverbal Social Behavior in Face to Face Interaction)

Funding SNSF – Ambizione

Coordinator Idiap Research Institute

Duration 01.11.2011 – 28.02.2015

Partner(s)

Name **MAAYA** (Multimedia Analysis and Access for Documentation and Decipherment of Maya Epigraphy)

Funding SNSF – DACH

Coordinator Idiap Research Institute

Duration 01.06.2013 – 31.05.2016

Partner(s) University of Geneva, University of Bonn

Name **SIVI** (Situating Vision to Perceive Object Shape and Affordances)

Funding SNSF – DACH

Coordinator University of Bonn

Duration 01.11.2011 – 30.11.2015

Partner(s) Technische Universität Wien, RWTH Aachen, Idiap Research Institute

Name **A-MUSE** (Adaptive Multilingual Speech Processing)

Funding FP7 – IP – ICT

Coordinator Fundacio Barcelona Media

Duration 01.12.2013 – 30.11.2016

Partner(s) Video Stream Networks, S.L., Université du Maine, Gottfried Wilhelm Leibniz Universität Hannover, Idiap Research Institute, Gesellschaft zur Förderung der angewandten Informationsforschung EV, Deutsche Welle

Name **DASH** (Object Detection with Active Sample Harvesting)

Funding SNSF – Division II

Coordinator Idiap Research Institute

Duration 01.09.2012 – 31.08.2015

Partner(s)

Name **UNITS** (Unified Speech Processing Framework for Trustworthy Speaker Recognition)

Funding SNSF – Division II

Coordinator Idiap Research Institute

Duration 01.07.2015 – 30.06.2018

Partner(s)

Name **G3E** (G3E: Geometric Generative Gaze Estimation model)

Funding SNSF – Division II

Coordinator Idiap Research Institute

Duration 01.04.2014 – 31.01.2015

Partner(s)

Name **HFACE** (Heterogeneous Face Recognition)

Funding SNSF – Division II

Coordinator Idiap Research Institute

Duration 01.07.2014 – 30.06.2017

Partner(s)

Name **PHASER** (PHASER: Parsimonious Hierarchical Automatic Speech Recognition)

Funding SNSF – Division II

Coordinator Idiap Research Institute

Duration 01.06.2014 – 31.05.2016

Partner(s)

ACTIVE AND GRANTED PROJECTS IN 2015

Name I-DRESS (Assistive Interactive robotic system for support in DRESSing)
Funding SNSF – ERA-NET Chist-Era
Coordinator Agencia Estatal Consejo Superior de Investigaciones Científicas
Duration 01.12.2015 – 30.11.2018
Partner(s) Idiap Research Institute, University of the West of England

Name YOUTH@NIGHT (Youth@Night – A multi-disciplinary multi-method study of young people's outgoing and drinking behaviors)
Funding SNSF – Interdisciplinary Project
Coordinator Sucht Schweiz – Research Institute
Duration 01.01.2014 – 31.12.2016
Partner(s) Universität Zürich, Idiap Research Institute

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

Name MODERN (Modeling discourse entities and relations for coherent machine translation)
Funding SNSF – Sinergia
Coordinator Fundacio Barcelona Media
Duration 01.12.2013 – 30.11.2016
Partner(s) Universiteit Utrecht, Universität Zürich, University of Geneva

Name SIWIS (Spoken Interaction with Interpretation in Switzerland)
Funding SNSF – Sinergia
Coordinator Idiap Research Institute
Duration 01.12.2012 – 30.11.2016
Partner(s) University of Geneva, University of Edinburgh, Eidgenössische Technische Hochschule Zürich

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

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

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

Name COHFACE (CONTACTless Heartbeat detection for trustworthy FACE Biometrics)
Funding Hasler Foundation
Coordinator Idiap Research Institute
Duration 01.09.2014 – 31.08.2015
Partner(s)

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

ACTIVE AND GRANTED PROJECTS IN 2015

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

Name **L-PASS** (Linguistic-Paralinguistic Speech Synthesis)
Funding Hasler Foundation
Coordinator Idiap Research Institute
Duration 01.12.2015 – 28.02.2016
Partner(s)

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

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

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

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

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

Name **RECOD2014** (low bit-rate speech coding)
Funding Armasuisse
Coordinator Idiap Research Institute
Duration 01.04.2014 – 31.03.2015
Partner(s)

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

Name **MIRROR** (Mobile Data to Improve Decision Making in African Cities)
Funding EPFL – Coopération
Coordinator Idiap Research Institute
Duration 01.01.2014 – 31.12.2016
Partner(s)

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

ACTIVE AND GRANTED PROJECTS IN 2015

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

Name VALAIS+ (Valais+ Une plateforme pour mieux connaître l'espace de vie du canton)
Funding Loterie Romande
Coordinator Idiap Research Institute
Duration 01.09.2015 – 30.08.2016
Partner(s) HES-SO Valais

PROJECTS FUNDED BY INDUSTRIAL PARTNERS

Name BIOWATCH (Biowatch)
Funding Fondation The Ark
Coordinator Idiap Research Institute
Duration 01.01.2015 – 30.04.2015
Partner(s) Biowatch SA

Name BAXTER (Démonstrateur Baxter)
Funding Fondation The Ark
Coordinator Idiap Research Institute
Duration 16.02.2015 – 15.04.2015
Partner(s)

Name EYEWARE (Eyeware)
Funding Fondation The Ark
Coordinator Idiap Research Institute
Duration 01.09.2015 – 30.04.2016
Partner(s)

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

Name HYBRID (Hybrid Recommender System)
Funding Fondation The Ark
Coordinator Unono
Duration 01.08.2015 – 31.03.2016
Partner(s) Idiap Research Institute

Name PUNK (Punktuation)
Funding Fondation The Ark
Coordinator recapp
Duration 01.05.2015 – 30.04.2016
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 01.03.2014 – 31.03.2015
Partner(s) Clinique romande de réadaptation, Cimark

Name BEEAPP (Beeapp)
Funding CTI
Coordinator Idiap Research Institute
Duration 01.10.2015 – 31.12.2015
Partner(s) Vergers des Essinges

Name DIGIT-ARENA (Real-Time Perimeter Board Content Digital Replacement)
Funding CTI
Coordinator Idiap Research Institute
Duration 01.12.2014 – 30.07.2016
Partner(s) E.S. Concept

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

ACTIVE AND GRANTED PROJECTS IN 2015

Name **MULTIVEO** (High Accuracy Speaker-Independent Multilingual Automatic Speech Recognition System)

Funding CTI

Coordinator Idiap Research Institute

Duration 01.11.2013 – 31.05.2015

Partner(s) Veovox

Name **SCOREL2** (Automatic scoring and adaptive pedagogy for oral language learning)

Funding CTI

Coordinator Idiap Research Institute

Duration 01.06.2014 – 31.01.2016

Partner(s) SpeedLingua

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 01.12.2015 – 31.07.2017

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

Name **VIDEOPROTECTOR** (Morphean VideoProtector)

Funding CTI

Coordinator Idiap Research Institute

Duration 01.06.2014 – 31.12.2015

Partner(s) HES-SO Fribourg

Name **ATAP** (Continuous Authentication Program)

Funding IND – Industrial Project

Coordinator Idiap Research Institute

Duration 15.01.2015 – 14.04.2015

Partner(s) Google

Name **DRACULA** (Detect and track people/object in order to deliver personalised movies.)

Funding IND – Industrial Project

Coordinator Idiap Research Institute

Duration 01.07.2015 – 29.02.2016

Partner(s) iRewind

Name **EMMA1** (Expression Mimics Marker Analysis)

Funding IND – Industrial Project

Coordinator Idiap Research Institute

Duration 01.03.2014 – 28.02.2015

Partner(s) DoctorWeb

Name **GOOGLE MOBILE** (Mobile Face and Voice Anti-Spoofing)

Funding IND – Industrial Project

Coordinator Idiap Research Institute

Duration 15.05.2015 – 31.03.2016

Partner(s)

PROJECTS AWARDED IN 2015 AND STARTING IN 2016

XXXIII

Name **MALORCA** (Machine Learning of Speech Recognition Models for Controller Assistance)

Funding H2020-SESAR-2015

Coordinator Deutsches Zentrum für Luft- und Raumfahrt EV

Duration 01.04.2016 – 31.03.2018

Partner(s) Idiap Research Institute, Austro Control Österreichische Gesellschaft für Zivilluftfahrt MBH, Řízení letového provozu České republiky, Universität des Saarlandes

Name **MUMMER** (MultiModal Mall Entertainment Robot)

Funding H2020-ICT-2015

Coordinator University of Glasgow

Duration 01.03.2016 – 28.02.2020

Partner(s) Kiinteistö Oy Ideapark AB, Centre national de la recherche scientifique, Teknologian Tutkimuskeskus VTT, ALDEBARAN Robotics, Idiap Research Institute

ACTIVE AND GRANTED PROJECTS IN 2015

Name **SUMMA** (Scalable Understanding of Multilingual Media)

Funding H2020-ICT-2015

Coordinator University of Edinburgh

Duration 01.02.2016 – 31.01.2019

Partner(s) LETA, University College London, Idiap Research Institute, Qatar Computing Research Institute, Deutsche Welle, Priberam Informatica SA, British Broadcast Corporation

Name **SWAN** (Secure Access Control over Wide Area Network)

Funding Research Council of Norway

Coordinator Hogskolen / Gjovik

Duration 01.01.2016 – 31.12.2018

Partner(s) Bankenverband, Universitaet Oslo, Morpho, Idiap Research Institute, Zwiipe AS

Name **TESLA** (An Adaptive Trust-based e-assessment System for Learning)

Funding H2020-ICT-2015

Coordinator Fundacio Per A La Universitat Oberta de Catalunya

Duration 01.01.2016 – 31.12.2018

Partner(s) Idiap Research Institute, The Open University, Anadolu University, Instituto Nacional de Astrofisica, Optica y Electronica, Technical University of Sofia, European Quality Assurance Network for Informatics Education e.V., Open Universiteit Nederland, WFSW, SA, Institut Mines-Telecom, Télécom SudParis, European Association for quality assurance in higher education AISBL, LPLUS GmbH, Sofiiski Universitet sveti kliment ohridski, Imperial College London, Protos Sistemas de Información, S.L., Agència per a la Qualitat del Sistema Universitari de Catalunya, Jyvaskylan Yliopisto, Université de Namur ASBL

Name **LIVEHEART** (The Cellular Basis of Cardiac Development Revealed by Live Imaging)

Funding SNSF – ANR

Coordinator Institut de génétique et de biologie moléculaire et cellulaire

Duration 01.03.2016 – 28.02.2019

Partner(s) Universität Bern, École Polytechnique Paris, Idiap Research Institute

Name **ODESSA** (Online Diarization Enhanced by recent Speaker identification and Sequential learning Approaches)

Funding SNSF – ANR

Coordinator Centre national de la recherche scientifique

Duration 01.03.2016 – 28.02.2019

Partner(s) EURECOM, Idiap Research Institute

Name **PLATFORM MMD** (Platform for Reproducible Acquisition, Processing, and Sharing of Dynamic, Multi-Modal Data)

Funding SNSF – R'Equip

Coordinator Idiap Research Institute

Duration 01.01.2016 – 31.12.2016

Partner(s)

Name **COMETS-M** (Computational Methods for Temporal Super-resolution Microscopy)

Funding SNSF – Division II

Coordinator Idiap Research Institute

Duration 01.04.2016–30.03.2019

Partner(s)

ACTIVE AND GRANTED PROJECTS IN 2015

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 01.04.2016 – 31.03.2019
Partner(s) Universitaet Bielfeld, Idiap Research Institute

Name **SMILE** (Scalable Multimodal sign language Technology for sign language Learning and assessmEnt)
Funding SNSF Sinergia
Coordinator Idiap Research Institute
Duration 01.03.2016 – 28.02.2019
Partner(s) University of Surrey, University of Applied Sciences of Special Needs Education

Name **ADDG2SU EXT.** (Flexible Acoustic data-driven Grapheme to Subword Unit Conversion)
Funding Hasler Foundation
Coordinator Idiap Research Institute
Duration 01.03.2016 – 28.02.2017
Partner(s)

Name **ELEARNING VALAIS 3.0** (eLearning-Valais 3.0)
Funding Fondation The Ark
Coordinator Formation universitaire à Distance
Duration 01.03.2016 – 31.01.2018
Partner(s) KLEWEL, Idiap Research Institute

Name **RECOMED2** (Intégration de la transcription vocale dans le dossier patient informatisé CRR)
Funding Fondation The Ark
Coordinator Idiap Research Institute
Duration 01.02.2016 – 30.06.2016
Partner(s) DPC Solutions, Clinique romande de réadaptation, recapp

Name **BIOWAVE** (A BIometric Watch Activated by VEins)
Funding CTI
Coordinator Centre Suisse d'Electronique et de Microtechnique
Duration 01.02.2016 – 31.07.2017
Partner(s) Idiap Research Institute, Biowatch SA

Name **ESGEM** (Enhanced Swiss German mEdia Monitoring)
Funding CTI
Coordinator Idiap Research Institute
Duration 01.03.2016 – 30.10.2017
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 01.05.2016 – 31.10.2017
Partner(s) KeyLemon SA

Books and Book Chapters

- [1] O. Aran, I. Ari, A. Kindiroglu, P. Santemiz, and L. Akarun, "Otomatik işaret dili tanıma ve türk işaret dili için bilgisayar uygulamaları," in *Türk İşaret Dili Araştırmaları - Research on Turkish Sign Language*, in Turkish, Koc University Press, 2015.
- [2] D. Bruno, S. Calinon, M. S. Malekzadeh, and D. G. Caldwell, "Learning the stiffness of a continuous soft manipulator from multiple demonstrations," in *Intelligent Robotics and Applications*, ser. Lecture Notes in Computer Science, H. Liu, N. Kubota, X. Zhu, R. Dillmann, and D. Zhou, Eds., vol. 9246, Best Paper Award Finalist at ICIRA'2015, Springer, 2015, pp. 185–195.
- [3] T. Kriyakerne and D. Ginsbourger, "Global optimization with sparse and local gaussian process models," in *Machine Learning, Optimization, and Big Data*, ser. Lecture Notes in Computer Science, P. Pardalos, M. Pavone, G. M. Farinella, and V. Cutello, Eds., vol. 9432, Springer International Publishing, 2015, pp. 185–196.
- [4] A. Lazaridis, B. Potard, and P. N. Garner, "DNN-based speech synthesis: Importance of input features and training data," in *International Conference on Speech and Computer*, SPECOM, ser. Lecture Notes in Computer Science, A. Ronzhin, R. Potapova, and N. Fakotakis, Eds., vol. 9319, Springer Berlin Heidelberg, 2015, pp. 193–200.
- [5] R. Lebret and R. Collobert, "Rehabilitation of count-based models for word vector representations," in *Computational Linguistics and Intelligent Text Processing*, ser. Lecture Notes in Computer Science, Alexander Gelbukh, vol. 9041, Springer International Publishing, 2015, pp. 417–429.
- [6] S. Marmin, C. Chevalier, and D. Ginsbourger, "Differentiating the multipoint expected improvement for optimal batch design," in *Machine Learning, Optimization, and Big Data*, ser. Lecture Notes in Computer Science, P. Pardalos, M. Pavone, G. M. Farinella, and V. Cutello, Eds., vol. 9432, Springer International Publishing, 2015, pp. 37–48.

- [7] B. Webber, M. Carpuat, A. Popescu-Belis, and C. Hardmeier, *Proceedings of the Second Workshop on Discourse in Machine Translation (DiscoMT 2015)*. Association for Computational Linguistics, Sep. 2015.

Articles in Scientific Journals

- [1] M. Cernak, P. N. Garner, A. Lazaridis, P. Motlicek, and X. Na, "Incremental syllable-context phonetic vocoding," *IEEE/ACM Transactions on audio, speech, and language processing*, vol. 23, no. 6, Jun. 2015.
- [2] I. Chingovska and A. Anjos, "On the use of client identity information for face anti-spoofing," *IEEE Transactions on Information Forensics and Security*, Special Issue on Biometric Anti-spoofing, vol. 10, no. 4, pp. 787–796, 2015.
- [3] A. Dutta, M. Günther, L. El Shafey, S. Marcel, R. Veldhuis, and L. Spreeuwers, "Impact of eye detection error on face recognition performance," *IET Biometrics*, vol. 4, no. 3, pp. 137–150, Sep. 2015.
- [4] 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*, pp. 1–23, Nov. 2015.
- [5] D. Gatica-Perez, "Signal processing in the workplace," *IEEE Signal Processing Magazine*, vol. 32, no. 1, pp. 121–125, Jan. 2015. *Vision*, pp. 1–23, Nov. 2015.
- [6] M. Guillemot, J.-M. Odobez, A. Vinciarelli, and S. Ingram, "KleweL webcast: From research to growing company," *IEEE Multimedia*, vol. 22, no. 4, pp. 94–99, Dec. 2015.
- [7] M. Günther, S. Böhringer, D. Wiczorek, and R. P. Würtz, "Reconstruction of images from Gabor graphs with applications in facial image processing," *Journal of Wavelets, Multiresolution and Information Processing*, vol. 13, no. 4, pp. 1550019 1–25, 2015.

LIST OF PUBLICATIONS IN 2015

- [8] M. Habibi and A. Popescu-Belis, "Keyword extraction and clustering for document recommendation in conversations," *IEEE/ACM Transactions on Audio Speech and Language Processing*, vol. 23, no. 4, pp. 746–759, 2015.
- [9] A. Hadid, N. Evans, S. Marcel, and J. Fierrez, "Biometrics systems under spoofing attack: An evaluation methodology and lessons learned," *IEEE Signal Processing Magazine*, vol. 32, no. 5, pp. 20–30, 2015.
- [10] R. Hu, G. Can, C. Pallan Gayol, G. Krempel, J. Spotak, G. Vail, S. Marchand-Maillet, J.-M. Odobez, and D. Gatica-Perez, "Multimedia analysis and access of ancient Maya epigraphy," *Signal processing magazine*, vol. 32, no. 4, pp. 75–84, 2015.
- [11] Y. E. Kara, G. Genc, O. Aran, and L. Akarun, "Modeling annotator behaviors for crowd labeling," *Neurocomputing*, vol. 160, no. C, pp. 141–156, Jul. 2015.
- [12] R. Lefort, L. Fusco, O. Pertz, and F. Fleuret, "Machine learning-based tools to model and to remove the off-target effect," *Pattern Analysis and Applications*, pp. 1–14, 2015.
- [13] T. Meyer, N. Hajlaoui, and A. Popescu-Belis, "Disambiguating discourse connectives for statistical machine translation," *IEEE/ACM Transactions on Audio, Speech and Language Processing*, vol. 23, no. 7, pp. 1184–1197, Jul. 2015.
- [14] P. Motlicek, D. Imseng, B. Potard, P. N. Garner, and I. Himawan, "Exploiting foreign resources for DNN-based ASR," *EURASIP Journal on Audio, Speech, and Music Processing*, no. 2015:17, pp. 1–10, Jun. 2015.
- [15] N. Pappas and A. Popescu-Belis, "Combining content with user preferences for non-fiction multimedia recommendation: A study on TED lectures," *Multimedia Tools and Applications, Special Issue on Content Based Multimedia Indexing*, pp. 1–23, Jan. 2015.
- [16] R. Rasipuram and M. Magimai-Doss, "Acoustic and lexical resource constrained ASR using language-independent acoustic model and language-dependent probabilistic lexical model," *Speech Communication*, vol. 68, pp. 23–40, Apr. 2015.
- [17] D. Sanchez-Cortes, S. Kumano, K. Otsuka, and D. Gatica-Perez, "In the mood for Vlog: Multimodal inference in conversational social video," *ACM Transactions on Interactive Intelligent Systems*, vol. 5, no. 2, 9:1–9:24, 2015.
- [18] A. Sapru and H. Bourlard, "Automatic recognition of emergent social roles in small group interactions," *IEEE Transactions on Multimedia*, vol. 17, no. 5, pp. 746–760, 2015.
- [19] B. Schuller, S. Steidl, A. Batliner, E. Nöth, A. Vinciarelli, F. Burkhardt, f. Weninger, F. Eyben, T. Bocklet, G. Mohammadi, and B. Weiss, "A survey on perceived speaker traits: Personality, likability, pathology and the first challenge," *Computer Speech and Language*, vol. 19, no. 1, pp. 100–131, Jan. 2015.
- [20] S. Sheikhi and J.-M. Odobez, "Combining dynamic head pose-gaze mapping with the robot conversational state for attention recognition in human-robot interactions," *Pattern Recognition Letters*, vol. 66, pp. 81–90, Nov. 2015.
- [21] A. Sizov, E. Khoury, T. Kinnunen, Z. Wu, and S. Marcel, "Joint speaker verification and antispoofing in the i-vector space," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 4, pp. 821–832, Apr. 2015.
- [22] N. Suditu and F. Fleuret, "Adaptive relevance feedback for large-scale image retrieval," *Multimedia wTools and Applications*, pp. 1–31, 2015.
- [23] M. J. Taghizadeh, A. Asaei, S. Haghighatshoar, P. N. Garner, and H. Bourlard, "Spatial sound localization via multipath euclidean distance matrix recovery," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 5, pp. 802–814, Aug. 2015.
- [24] V. Trivedi, T. V. Truong, L. A. Trinh, D. B. Holland, M. Liebling, and S. E. Fraser, "Dynamic structure and protein expression of the live embryonic heart captured by 2-photon light sheet microscopy and retrospective registration," *Biomedical Optics Express*, vol. 6, no. 6, pp. 2056–2066, Jun. 2015.

- [25]** 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*, in press.
- [26]** A. Asaei, M. J. Taghizadeh, S. Haghghatshoar, 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*, in press.
- [27]** 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*, in press.
- [28]** N. Pappas and A. Popescu-Belis, "Adaptive sentiment-aware one-class collaborative filtering," *Expert Systems with Applications*, in press.
- [29]** O. Pertz, L. Fusco, R. Lefort, K. C. Smith, F. Benmansour, G. Gonzalez, C. Barilari, B. Rinn, F. Fleuret, and P. Fua, "Computer vision profiling of neurite outgrowth dynamics reveals spatiotemporal modularity of Rho GTPase signaling," *Journal of Cell Biology*, in press.
- [30]** R. Rasipuram and M. Magimai-Doss, "Articulatory feature based continuous speech recognition using probabilistic lexical modeling," *Computer Speech and Language*, in press.
- [31]** R. Ullmann and H. Bourlard, "Predicting the intrusiveness of noise through sparse coding with auditory kernels," *Speech Communication*, in press.

Theses

- [1]** I. Chingovska, "Trustworthy biometric verification under spoofing attacks: Application to the face mode," PhD thesis, École Polytechnique Fédérale de Lausanne, Nov. 2015.
- [2]** K. A. Funes Mora, "3D gaze estimation from remote RGB-D sensors," PhD thesis, École Polytechnique Fédérale de Lausanne, Oct. 2015.
- [3]** M. Habibi, "Modeling users' information needs in a document recommender for meetings," PhD thesis, École Polytechnique Fédérale de Lausanne, Nov. 2015.
- [4]** L. S. Nguyen, "Computational analysis of behavior in employment interviews and video resumes," PhD thesis, École Polytechnique Fédérale de Lausanne, May 2015.
- [5]** A. Sapru, "Automatic social role recognition and its application in structuring multiparty interactions," PhD thesis, École Polytechnique Fédérale de Lausanne, Jun. 2015.
- [6]** M. J. Taghizadeh, "Enabling speech applications using ad-hoc microphone arrays," PhD thesis, École Polytechnique Fédérale de Lausanne, May 2015.
- [7]** S. H. Yella, "Speaker diarization of spontaneous meeting room conversations," PhD thesis, École Polytechnique Fédérale de Lausanne, Lausanne, Jan. 2015.

Articles in Conference Proceedings

- [1] A. Asaei, M. Cernak, and H. Boulard, "On compressibility of neural network phonological features for low bit rate speech coding," in *Proceeding of Interspeech*, 2015.
- [2] A. Asaei, N. Mohammadiha, M. J. Taghizadeh, S. Doclo, and H. Boulard, "On application of non-negative matrix factorization for ad hoc microphone array calibration from incomplete noisy distances," in *IEEE 40th International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2015.
- [3] T. Bagautdinov, F. Fleuret, and P. Fua, "Probability occupancy maps for occluded depth images," in *Proceedings of the IEEE international conference on Computer Vision and Pattern Recognition*, 2015.
- [4] S. Calinon, "Robot learning with task-parameterized generative models," in *Proc. Intl Symp. on Robotics Research*, 2015.
- [5] M. Cernak and P.-E. Honnet, "An empirical model of emphatic word detection," in *Proc. of Interspeech*, Dresden, Germany: ISCA, Sep. 2015.
- [6] M. Cernak, B. Potard, and P. N. Garner, "Phonological vocoding using artificial neural networks," in *IEEE 40th International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brisbane, Australia: IEEE, Apr. 2015.
- [7] N. Chacko, K. G. Chan, and M. Liebling, "Intensity-based point-spread-function-aware registration for multi-view applications in optical microscopy," in *Biomedical Imaging (ISBI)*, 2015 IEEE 12th International Symposium on, Apr. 2015, pp. 306–309.
- [8] K. G. Chan and M. Liebling, "Estimation of divergence-free 3D cardiac blood flow in a zebrafish larva using multi-view microscopy," in *Biomedical Imaging (ISBI)*, 2015 IEEE 12th International Symposium on, IEEE, Brooklyn, NY, USA, Apr. 2015, pp. 385–388.
- [9] G. Chávez-Martínez, S. Ruiz-Correa, and D. Gatica-Perez, "International conference on mobile and ubiquitous multimedia," in *Happy and Agreeable? Multi-Label Classification of Impressions in Social Video*, Linz, Dec. 2015.
- [10] Y. Chen, Y. Yu, and J.-M. Odobez, "Head nod detection from a full 3D model," in *Proceedings of the ICCV 2015*, Dec. 2015.
- [11] P. Dighe, A. Asaei, and H. Boulard, "Dictionary learning for sparse representation of neural network exemplars in speech recognition," in *Proceedings of SPARS 2015: Workshop on Signal Processing with Adaptive Sparse Structured Representations*, 2015, 2015.
- [12] P. Dighe, A. Asaei, and H. Boulard, "Sparse modeling of neural network posterior probabilities for exemplar-based speech recognition," in *Proceedings of SPARS 2015: Workshop on Signal Processing with Adaptive Sparse Structured Representations*, 2015.
- [13] N. Erdogmus, M. Vanoni, and S. Marcel, "Within- and cross- database evaluations for gender classification via BeFIT protocols," in *International Workshop on Multimedia Signal Processing*, 2015.
- [14] T. de Freitas Pereira and S. Marcel, "Periocular biometrics in mobile environment," in *IEEE Seventh International Conference on Biometrics: Theory, Applications and Systems*, Arlington, USA: IEEE, Sep. 2015.
- [15] J. Gancet, D. Urbina, P. Letier, M. Ilzokvitz, P. Weiss, F. Gauch, G. Antonelli, G. Indiveri, G. Casalino, A. Birk, M. F. Pfungsthorn, S. Calinon, A. Tanwani, A. Turetta, C. Walen, and L. Guilpain, "Dexrov: Dexterous undersea inspection and maintenance in presence of communication latencies," in *IFAC Workshop on Navigation, Guidance and Control of Underwater Vehicles (NGCUV)*, 2015.
- [16] B. Gerazov and P. N. Garner, "An investigation of muscle models for physiologically based intonation modeling," in *Proceedings of the 23rd Telecommunications Forum*, Belgrade, Serbia, Nov. 2015.
- [17] B. Gerazov, P.-E. Honnet, A. Gjoreski, and P. N. Garner, "Weighted correlation based atom decomposition

LIST OF PUBLICATIONS IN 2015

intonation modelling," in Proceedings of Interspeech, Sep. 2015.

- [18]** G. Gosztolya, T. Grósz, L. Tóth, and D. Imseng, "Building context-dependent DNN acoustic models using Kullback-Leibler divergence-based state tying," in Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing, 2015.
- [19]** M. Habibi and A. Popescu-Belis, "Query refinement using conversational context: A method and an evaluation resource," in Proceedings of the NLDB 2015 (20th International Conference on Applications of Natural Language to Information Systems), 2015.
- [20]** I. Himawan, P. Motlicek, M. Ferras, and S. Madikeri, "Towards utterance-based neural network adaptation in acoustic modeling," in IEEE Automatic Speech Recognition and Understanding Workshop, 2015.
- [21]** I. Himawan, P. Motlicek, D. Imseng, B. Potard, N. Kim, and J. Lee, "Learning feature mapping using deep neural network bottleneck features for distant large vocabulary speech recognition," in IEEE International Conference on Acoustics, Speech, and Signal Processing, 2015.
- [22]** I. Himawan, P. Motlicek, S. Sridharan, D. Dean, and D. Tjondronegoro, "Channel selection in the short-time modulation domain for distant speech recognition," in Proceedings of Interspeech, 2015.
- [23]** P.-E. Honnet, B. Gerazov, and P. N. Garner, "Atom decomposition-based intonation modelling," in IEEE 40th International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brisbane, Australia: IEEE, Apr. 2015.
- [24]** A. Hyafil and M. Cernak, "Neuromorphic based oscillatory device for incremental syllable boundary detection," in Proc. of Interspeech, Dresden, Germany, Sep. 2015.
- [25]** B. Jou, T. Chen, N. Pappas, M. Redi, M. Topkara, and S.-F. Chang, "Visual affect around the world: A large-scale multilingual visual sentiment ontology," in ACM International Conference on Multimedia, 2015.
- [26]** E. Khan, P. Baqué, F. Fleuret, and P. Fua, "Kullback-Leibler proximal variational inference," in Proceedings

of the international conference on Neural Information Processing Systems, 2015.

- [27]** S. Kucur Ergunay, E. Khoury, A. Lazaridis, and S. Marcel, "On the vulnerability of speaker verification to realistic voice spoofing," in IEEE International Conference on Biometrics: Theory, Applications and Systems, Sep. 2015.
- [28]** I. Kuzborskij, F. Orabona, and B. Caputo, "Transfer learning through greedy subset selection," in Volume 9279 of the series Lecture Notes in Computer Science, 2015.
- [29]** N. Le, D. Wu, S. Meignier, and J.-M. Odobez, "EUMSSI team at the mediaeval person discovery challenge," in Working Notes Proceedings of the MediaEval 2015 Workshop, Wurzen, Germany, Sep. 2015.
- [30]** R. Lebrete and R. Collobert, "N-gram-based low-dimensional representation for document classification," in International Conference on Learning Representations, 2015.
- [31]** R. Lebrete, P. H. O. Pinheiro, and R. Collobert, "Phrase-based image captioning," in International Conference on Machine Learning (ICML), 2015.
- [32]** J. Legrand and R. Collobert, "Joint RNN-based greedy parsing and word composition," in Proceedings of ICLR 2015, 2015.
- [33]** N.-Q. Luong, L. M. Werlen, and A. Popescu-Belis, "Pronoun translation and prediction with or without coreference links," in EMNLP 2015 Workshop on Discourse in Machine Translation, Sep. 2015.
- [34]** S. Madikeri and H. Bourlard, "KL-HMM based speaker diarization system for meetings," in Proceedings of ICASSP 2015, Apr. 2015, pp. 4435-4439.
- [35]** S. Madikeri, I. Himawan, P. Motlicek, and M. Ferras, "Integrating online l-vector extractor with information bottleneck based speaker diarization system," in Proceedings of Interspeech 2015, 2015.
- [36]** S. Madikeri, P. Motlicek, and H. Bourlard, "Combining SGMM speaker vectors and KL-HMM approach for

LIST OF PUBLICATIONS IN 2015

- speaker diarization," in Proceedings of ICASSP 2015, 2015, pp. 4834-4837.
- [37]** R. Metha, M. Günther, and S. Marcel, "Gender classification by LUT based boosting of overlapping block patterns," in Scandinavian Conference on Image Analysis, vol. 9127, Springer International Publishing, 2015, pp. 530-542.
- [38]** P. Motlicek, S. Dey, S. Madikeri, and L. Burget, "Employment of subspace gaussian mixture models in speaker recognition," in 2015 IEEE International Conference on Acoustics, Speech, and Signal Processing, IEEE, Brisbane, Australia, May 2015, pp. 4445-4449.
- [39]** L. S. Nguyen and D. Gatica-Perez, "I would hire you in a minute: Thin slices of nonverbal behavior in job interviews," in Proceedings of the ACM International Conference on Multimodal Interaction, 2015.
- [40]** C. Oertel, K. A. Funes Mora, J. Gustafson, and J.-M. Odobez, "Deciphering the silent participant. On the use of audio-visual cues for the classification of listener categories in group discussions," in International Conference on Multimodal Interaction, Seattle, Washington, USA: ACM, Nov. 2015.
- [41]** S. Okada, O. Aran, and D. Gatica-Perez, "Personality trait classification via co-occurrent multiparty multimodal event discovery," in Proceedings of the ACM International Conference on Multimodal Interaction, 2015.
- [42]** D. Palaz, M. Magimai.-Doss, and R. Collobert, "Analysis of CNN-based speech recognition system using raw speech as input," in Proceedings of Interspeech, 2015.
- [43]** D. Palaz, M. Magimai.-Doss, and R. Collobert, "Convolutional neural networks-based continuous speech recognition using raw speech signal," in International Conference on Acoustics, Speech and Signal Processing, Apr. 2015.
- [44]** P. H. O. Pinheiro and R. Collobert, "From image-level to pixel-level labeling with convolutional networks," in Computer Vision and Pattern Recognition (CVPR), 2015.
- [45]** P. H. O. Pinheiro, R. Collobert, and P. Dollár, "Learning to segment objects candidates," in Advances in Neural Information Processing Systems, 2015.
- [46]** X. Pu, L. Mascarell, A. Popescu-Belis, M. Fishel, N.-Q. Luong, and M. Volk, "Leveraging compounds to improve noun phrase translation from chinese and german," in Proceedings of the ACL-IJCNLP 2015 Student Research Workshop, Jul. 2015.
- [47]** R. Raghavendra, M. Avinas, C. Busch, and S. Marcel, "Finger vein liveness detection using motion magnification," in IEEE International Conference on Biometrics: Theory, Applications and Systems, Sep. 2015.
- [48]** D. Ram, A. Asaei, P. Dighe, and H. Bourlard, "Sparse modeling of posterior exemplars for keyword detection," in Proceedings of Interspeech, 2015.
- [49]** R. Rasipuram, M. Cernak, A. Nanchen, and M. Magimai.-Doss, "Automatic accentedness evaluation of non-native speech using phonetic and sub-phonetic posterior probabilities," in Proceedings of Interspeech, Sep. 2015.
- [50]** R. Rasipuram, M. Razavi, and M. Magimai.-Doss, "Integrated pronunciation learning for automatic speech recognition using probabilistic lexical modeling," in International Conference on Acoustics, Speech and Signal Processing, 2015.
- [51]** M. Razavi and M. Magimai.-Doss, "An HMM-based formalism for automatic subword unit derivation and pronunciation generation," in International Conference on Acoustics, Speech and Signal Processing, 2015.
- [52]** M. Razavi, R. Rasipuram, and M. Magimai.-Doss, "Pronunciation lexicon development for under-resourced languages using automatically derived subword units: A case study on scottish gaelic," in 4th Biennial Workshop on Less-Resourced Languages, 2015.
- [53]** L. Rozo, D. Bruno, S. Calinon, and D. G. Caldwell, "Learning optimal controllers in human-robot

LIST OF PUBLICATIONS IN 2015

cooperative transportation tasks with position and force constraints," in Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS), Hamburg, Germany, 2015.

[54] D. Santani and D. Gatica-Perez, "Loud and trendy: Crowdsourcing impressions of social ambiance in popular indoor urban places," in Proceedings of the 23rd ACM International Conference on Multimedia, ser. MM '15, Brisbane, Australia: ACM, Oct. 2015, pp. 211–220.

[55] D. Santani, J. Njuguna, T. Bills, A. W. Bryant, R. Bryant, J. Ledgard, and D. Gatica-Perez, "CommuniSense: Crowdsourcing road hazards in Nairobi," in Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services, ser. MobileHCI '15, Copenhagen, Denmark: ACM, Aug. 2015, pp. 445–456.

[56] D. Santani, S. Ruiz-Correa, and D. Gatica-Perez, "Looking at cities in Mexico with crowds," in Proceedings of the 2015 Annual Symposium on Computing for Development, ser. DEV '15, London, United Kingdom: ACM, Dec. 2015, pp. 127–135.

[57] B. Schuller, B. Vlasenko, F. Eyben, M. Wöllmer, A. Stuhlsatz, A. Wendemuth, and G. Rigoll, "Cross-corpus acoustic emotion recognition: Variances and strategies (extended abstract)," in Proceedings of ACII, 2015.

[58] I. Siegert, R. Boeck, B. Vlasenko, and A. Wendemuth, "Exploring dataset similarities using PCAbased feature selection," in Proceedings of ACII, 2015.

[59] I. Siegert, R. Boeck, B. Vlasenko, K. Ohnemus, and A. Wendemuth, "Overlapping speech, utterance duration and affective content in HHI and HCI – an comparison," in 6th IEEE Conference on Cognitive Infocommunications, 2015.

[60] J. Silverio, L. Rozo, S. Calinon, and D. G. Caldwell, "Learning bimanual end-effector poses from demonstrations using task-parameterized dynamical systems," in Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS), Hamburg, Germany, 2015.

[61] M. J. Taghizadeh, S. Haghghatshoar, A. Asaei, P. N.

Garner, and H. Boulard, "Robust microphone placement for source localization from noisy distance measurements," in IEEE 40th International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2015.

[62] P. Tome and S. Marcel, "On the vulnerability of palm vein recognition to spoofing attacks," in The 8th IAPR International Conference on Biometrics (ICB), May 2015.

[63] P. Tome and S. Marcel, "Palm vein database and experimental framework for reproducible research," in IEEE International Conference of the Biometrics Special Interest Group, Sep. 2015.

[64] P. Tome, R. Raghavendra, C. Busch, S. Tirunagari, N. Poh, B. H. Shekar, D. Gragnaniello, C. Sansone, L. Verdoliva, and S. Marcel, "The 1st competition on counter measures to finger vein spoofing attacks," in The 8th IAPR International Conference on Biometrics (ICB), May 2015.

[65] T. Tommasi, N. Patricia, B. Caputo, and T. Tuytelaars, "A deeper look at dataset bias," in German Conference on Pattern Recognition, 2015.

[66] R. Ullmann, M. Magimai-Doss, and H. Boulard, "Objective speech intelligibility assessment through comparison of phoneme class conditional probability sequences," in 40th IEEE Int. Conf. on Acoustics, Speech and Signal Processing, Brisbane, Australia, 2015.

[67] R. Ullmann, R. Rasipuram, M. Magimai-Doss, and H. Boulard, "Objective intelligibility assessment of text-to-speech systems through utterance verification," in Proceedings of Interspeech, 2015.

[68] J. Velasco, M. J. Taghizadeh, A. Asaei, H. Boulard, C. J. Martín-Arguedas, J. Macias-Guarasa, and D. Pizarro, "Novel GCC-PHAT model in diffuse sound field for microphone array pairwise distance based calibration," in IEEE 40th International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2015.

[69] B. Vlasenko and A. Wendemuth, "Annotators'

LIST OF PUBLICATIONS IN 2015

agreement and spontaneous emotion classification performance," in Proceedings of Interspeech, 2015.

- [70]** M. Zeestraten, S. Calinon, and D. G. Caldwell, "Learned minimal intervention control synthesis based on hidden semi-Markov models," in Proc. of the 8th Intl Workshop on Human-Friendly Robotics, Munich, Germany, Oct. 2015, p. 17.

I — XLIV

encarté scientifique