Séminaires S2A mars et avril 2018

Il y aura plusieurs séminaires en mars et en avril 2018 :

- Mercredi 21 mars à 16h en C47 — **Claude Barras**

  Titre:
  Reconnaissance du locuteur et structuration en tours de parole – travaux récents au LIMSI

  Résumé:
  Je présenterai des travaux actuellement en cours au LIMSI dans le domaine de la reconnaissance du locuteur et de la structuration en tours de parole ("speaker diarization"), menés en particulier dans le cadre du projet ANR-FNS ODESSA (Online Diarization Enhanced by recent Speaker identification and Structured prediction Approaches). Il s'agit d'un projet bilatéral franco-suisse conduit par le LIMSI en collaboration avec Eurecom et l'Idiap. Bien que la structuration en tours de parole ait fait l'objet de campagnes d'évaluation internationales depuis une quinzaine d'années, il ne s'agit cependant pas d'un problème résolu et j'en expliquerai quelques limites toujours actuelles. Avec un doctorant, nous avons récemment montré l'apport de réseaux neuronaux récurrents (LSTM) pour détecter les changements de locuteurs et pour effectuer une re-segmentation temporelle précise du signal. Nous avons aussi proposé un protocole de repérage à faible latence des locuteurs dans un flux acoustique, afin de prendre rapidement la meilleure décision possible.

  Bio:
  Claude Barras est maître de conférences en informatique à l'Université Paris-Sud. Après des études d'ingénieur à Supélec et une thèse soutenue en 1996 à l'Université Pierre-et-Marie-Curie sur la reconnaissance de parole, il rejoint le groupe Traitement du Langage Parlé du LIMSI en 2000. Ses travaux portent sur la structuration et la transcription enrichie de documents multimédia et sur la reconnaissance du locuteur. Il a été coordinateur de plusieurs projets ANR et a publié environ 80 articles en conférences et journaux à comité de lecture. Il est habilité à diriger des recherches et co-encadre plusieurs doctorants. Ses enseignements couvrent les différents domaine de l'informatique et plus particulièrement l'apprentissage automatique et le traitement de la parole.

- Jeudi 22 mars à 14h en amphi Saphir — **Laurent Daudet**

  Title: "From computational imaging to optical computing"

  Abstract:
  It has long been considered that the multiple diffusion resulting from the propagation of waves in disordered environments, such as light through biological tissues or a fine layer of paint, destroys all the information carried by these waves. In recent years, particularly in optics with wavefront control techniques, it has been shown that such propagation is extremely complex but remains linear. In fact, with spatially discrete inputs and outputs, it performs the equivalent of a random projection, i.e. the multiplication of the input vector by an iid random matrix. In this context, we have shown that these environments act precisely as "compressed sensing" model systems, allowing signal acquisition with a number of measurements driven by the actual amount of information. Conversely, one can see this physical system as an optimal mixer of information, performing instantaneously in the (physical) analog domain an elementary computation brick of many Machine Learning schemes. We will present a series of proof-of-concept experiments of image classification and transfer learning, based on these random features, and discuss recent technological developments of such optical co-processors.

- Mardi 27 mars à 10h en amphi Jade — **Maria Vakalopoulou** (Center for Visual Computing laboratory, CentraleSupélec)

  Title: Advanced Computer Vision and Machine Learning Techniques for Remote Sensing and Medical Applications.

Abstract:
Recent advancements in the computer vision and machine learning communities have drawn the attention of a big range of scientific communities as they have reached excellent performance on a variety of classification and detection problems. In this talk, I will present part of my work that focuses on the development of advanced machine learning and vision techniques and their application on medical imagery and remote sensing. In particular, in the first part of my presentation I will discuss the problems of image registration, semantic segmentation and change detection for optical very high resolution remote sensing images and I will focus on a novel supervised algorithm which addresses these three problems simultaneously in one joint framework using graphical models and higher order dependences. In the second part, I will discuss the application of deep learning methods in medical imagery. Namely, I will describe a novel method that alleviates the problem of limited training data by projecting the subjects to different anatomical atlases. We show that applying deep learning on this augmented dataset provides more accurate segmentation of the scleroderma lung disease.

Short bio:
Maria Vakalopoulou received an Engineering Diploma degree in survey engineering, graduating with excellence from the National Technical University of Athens, Athens, Greece, in 2011, and her PhD, in 2017, from the same university. During 2014 and 2015, she was a Visiting Student at University Paris-Est, École des Ponts ParisTech under the supervision of Prof. N. Komodakis and Prof. N. Paragios. Currently, she is a postdoctoral researcher at Center for Visual Computing laboratory of CentraleSupélec, University Paris-Saclay, Paris, France, working with Prof. N. Paragios. Her research interests include remote sensing, medical imagery, computer vision, and machine learning.

- Mercredi 28 mars à 10h en C48 — **Hendrik Purwins**

  Title:
  Cognitively Plausible Machine Learning for Audio Signal Processing

  Abstract:

  First, I will present (biologically inspired) sparse approximation methods for sound classification and synthesis. In Scholler & Purwins (JSTSP 2011), the shapes and lengths of the atoms of a sound dictionary are learned and a spike representation of the signal is used for sound classification. Clustering sound atoms from a sparse representation as a foreground layer and modeling the background noise by an LPC-based model leads to a method for sound texture re-synthesis (Kersten & Purwins, 2010,2011, 2012, http://tinyurl.com/jbdsvov).
  Secondly, I will talk about a statistical music representation based on cognitive principles (unsupervised, life-long, and one-shot learning, cognitive-perceptual top-down control) that are scalable in complexity. Applications are given for the generation of stylistically similar and musically interesting variations of a given piece of audio. (see http://tinyurl.com/zpx22mo, Marchini & Purwins 2011; Marxer & Purwins IEEE TASP, 2016).
  Third, I will introduce the Musical Brain Computer Interface and present found neural correlates of selective attention to voices in polyphonic music, based on ERP analysis in a multi-streamed oddball experiment (Treder, Purwins et al., JNE 2014).
  I will also briefly mention on-going work on deep learning for binary mask estimation in source separation and reinforcement learning for training agents in computer games.
  Finally I will outline my research vision of modelling the loop of musical creation and learning, combining various machine learning paradigms.

- Mercredi 4 avril à 14h en B555 — **Soledad Villar**

  Title:
  K-means clustering with optimization

  Abstract:
  K-means clustering aims to partition a set of n points into k clusters in such a way that each observation belongs to the cluster with the nearest mean, and such that the sum of squared distances from each point to its nearest mean is minimal. In the worst case, this is a hard optimization problem, requiring an exhaustive search over all possible partitions of the data into k clusters in order to find

the optimal clustering. At the same time, fast heuristic algorithms for k-means are widely used for data science applications, despite only being guaranteed to converge to local minimizers of the k-means objective.

In this talk, we consider a semidefinite programming relaxation of the k-means optimization problem. We discuss two regimes where the SDP provides an algorithm with improved clustering guarantees compared to previous results in the literature: (a) for points drawn from isotropic distributions supported in separated balls, the SDP recovers the globally optimal k-means clustering under mild separation conditions; (b) for points drawn from mixtures of distributions with bounded variance, the SDP solution can be rounded to a clustering which is guaranteed to classify all but a small fraction of the points correctly.

An interesting feature about the theoretical tools developed for proving (approximate) optimality of partitions under models (a) and (b) is that they can also be used to a posteriori certify (approximate) optimality of k-means clustering solutions of real data, no model required.

- Jeudi 5 avril à 14h en B316 — **Mathieu Lerasle** (CNRS, Psud)

  Title:
  MOM pour l'apprentissage robuste

  Abstract:

  Je présenterai les estimateurs par median-of-means et étudierai leurs propriétés de concentration. Muni de ces résultats, je montrerai comment construire des estimateurs atteignant des vitesses sous-gaussiennes dans certains problèmes d'apprentissage simples.
  Si le temps le permet, je montrerai également comment démontrer ces résultats en utilisant un parallèle avec la théorie des tests multiples.
  Je donnerai le principe de construction d'algorithmes basé sur le principe MOM et en discuterai certains avantages et inconvénients pratiques.

- Jeudi 5 avril à 16h en C48 — **Geoffroy Peeters**

  Title:
  "Recherche récente en analyse, séparation et synthèse musicale"

  Abstract:

  Lors de ce séminaire, je parlerai de nos recherches récentes dans le domaine de la séparation de source, de l'accroissement des bases d'entraînement (par paradigme teacher-student ou par data augmentation) ainsi que de l'utilisation de techniques de factorisation matricielle pour la synthèse audio (audio mosaicing).

  Bio:

  Geoffroy Peeters is a senior-researcher at IRCAM where he leads research activities related to music information retrieval from audio.
  He received his PHDs degree in 2001 and Habilitation in 2013 from University Paris-VI on audio signal processing, data analysis and machine learning. He has developed new algorithms to describe timbre, automatic classification, audio identification, rhythm description, musical structure and audio summary generation. He is the author of numerous articles and several patents in these areas. He is co-author of the ISO MPEG-7 audio standard. He has been co-general chair of the LSAS-2008, DAFx-2011 and will be of the ISMIR-2018 conference. He is member of the DAFx board, IEEE Task Force on Computational Audio Processing and member elected of the ISMIR board for the period 2016-2017 and 2018-2019.

- Jeudi 16 avril 16h — **Nicolas Keriven**

  Title: Sketched Learning from Random Features Moments

Abstract: Learning parameters from voluminous data can be prohibitive in terms of memory and computational requirements. Furthermore, modern architectures often ask for learning methods to be amenable to streaming or distributed computing. In this context, a popular approach is to first compress the database into a representation called a linear sketch, then learn the desired information using only this sketch. In this talk, we introduce a methodology to fit a mixture of probability distributions on the data, using only a sketch of the database. The sketch is defined by combining two notions from the reproducing kernel literature, kernel mean embedding and random features. It is seen to correspond to linear measurements of the probability distribution of the data, and the problem is thus analyzed under the lens of Compressive Sensing (CS), in which a signal is randomly measured and recovered. We analyze the problem using two classical approaches in CS: first a Restricted Isometry Property in the Banach space of finite signed measures, from which we obtain strong recovery guarantees however with an intractable non-convex minimization problem, and second with a dual certificate analysis, from which we show that total-variation regularization yields a convex minimization problem that in some cases recovers exactly the number of components of a gaussian mixture model. We also briefly describe a flexible heuristic greedy algorithm to estimate mixture models from a sketch, and apply it on synthetic and real data.

- Jeudi 16 avril 17h — **Guillaume Garrigos**

Title: Structured sparsity in inverse problems and support recovery with mirror-stratifiable functions
Abstract: We consider inverse problems where the prior on the data is an assumption of structured sparsity, and we look at a class of regularizers for which minimization algorithms identify in finite time some extended support of the original data. This is a direct consequence of a more general identification theorem, involving the so-called mirror stratifiability of the regularizer, a notion developped recently, and based on duality arguments.

We provide necessary and sufficient conditions for norm regularizers to be mirror stratifiable, and show its tight relationship with the geometry of the corresponding unit ball. Then we explain why exact model consistency cannot hold in general, but how we can nevertheless manage to ensure an enlarged model, by means of a dual certificate.

We discuss in particular whether stochastic algorithms can (or cannot) enjoy this identification property, in a statistical learning context.

We finally discuss how this approach can be extended to problems in separable Hilbert spaces, (Multiple Kernel Learning for instance). As a by-product, we derive improved rates of convergence for the minimization algorithms, like a new linear rate result for the iterative soft-thresholding algorithm in L2, with no assumptions.

Works in collaboration with:
J. Fadili, J. Malick, G. Peyré, L. Rosasco, S. Villa

ArXiv references:
1803.08381 , 1803.00783 and 1712.00357