

INTERNSHIP PROPOSAL

IFP Energies nouvelles

Subject

Compatibility between sparse machine learning and lossy compression

Profile

Master Student in Electrical and Computer Engineering or Master, Computer Science/Algorithmics, Image and Signal Processing, Statistics, Machine Learning, Applied Mathematics

Project Context

Many experimental designs acquire continuous or salve signals or images. Those are characteristic of a specific phenomenon. One may find examples at IFPEN in seismic data/images, NDT/NDE acoustic emissions (corrosion, battery diagnosis) engine benches (cylinder pressure data, fast camera), high-throughput screening in chemistry. Very often, such data is analyzed with standardized, a priori indices. Comparisons between different experiments (difference- or classification-based) are often based on the same indices, without resorting to initial measurements.

The increasing data volume, the variability in sensor and sampling, the possibility of different pre-processing yield two problems: the management and access to data (« *big data* ») and their optimal exploitation by dimension reduction methods, supervised or unsupervised learning (« *data science* »). This project aims at the analysis of the possibility of a joint compressed representation of data and the extraction of pertinent indicators, at different characteristic scales, and the relative impact of the first aspect (lossy compression degradation) over the second aspect (precision and robustness of extracted feature indicators).

Internship tutors

Camille Couprie/Laurent Duval (IFPEN)

Intership objectives

The internship possesses a dual goal. The first aspect will be dealing with scientific research on sparse signal/image representations with convolution networks based on multiscale wavelet techniques, called scattering networks. Their descriptors (or footprints) possess fine translation, rotation and scale invariance. Those descriptors will be employed for classification and detection. The second aspect will carry on the impact of lossy compression on the preceeding results, and the development of novel sparse representations for joint compression and learning.

J. Bruna, S. Mallat. [Invariant scattering convolution networks](#). *IEEE Trans. on Patt. Anal. and Mach. Int.*, 2010

L. Jacques, L. Duval, C. Chaux, G. Peyré, [A Panorama on Multiscale Geometric Representations, Intertwining Spatial, Directional and Frequency Selectivity](#), Signal Processing, 2011

C. Couprie, C. Farabet, L. Najman, Y. LeCun, [Convolutional Nets and Watershed Cuts for Real-Time Semantic Labeling of RGBD Videos](#), Journal of Machine Learning Research, 2014

A PhD thesis (Characteristic fingerprint computation and storage for high-throughput data flows and their on-line analysis, with J.-C. Pesquet, Univ. Paris-Est) is proposed, starting September 2015.

Information update: <http://www.laurent-duval.eu/lcd-2015-intern-learning-compression.html>

Duration: 5 to 6 months

Year: 2015

Location : IFPEN – Rueil (F-92000)

Accomodation : ~900 to 1000 euros/month dep. education

Candidature (CV & motivations): Camille Couprie / Laurent Duval

IFPEN Direction Mécatronique et numérique

Phone : +33(0)147525082/+33(0)147526102

e-mail : [camille.couprie](mailto:camille.couprie@laurent-duval.fr) / laurent.duval [at] ifpen.fr