Journée Recherche IDS

6 Juillet 2017

Hotel des Arts et Métiers
9 bis, avenue d'Iéna - 75016 PARIS

Programme

08h30 -- 09h05 : Accueil, café
09h05 -- 09h50 : *A guided tour in Monte Carlo*, F. Portier
09h50 -- 10h10 : *Parsimonious representations for modeling, visualizing and statistically analyzing brain tractograms*, P. Gori
10h10 -- 10h30 : *Restarting accelerated gradient methods with a rough strong convexity estimate*, O. Fercoq

10h30 -- 10h50 : Café

10h50 -- 11h10 : *User-centered modeling of skeletal weights for efficient 3D character animation*, J.-M. Thierry
11h10 -- 11h30 : *Study of User Engagement in Spontaneous Human-Robot Interactions: Case of User Engagement Breakdown*, A. Ben Youssef
11h30 -- 12h00 : Présentation des activités récentes des groupes, Responsables de groupes

12h00 -- 13h30 : Déjeuner

13h30 -- 14h15 : *L'esthétique dans les images*, H. Maître
14h15 -- 14h35 : *Faster Independent Component Analysis (ICA) by preconditionning with Hessian approximations*, A. Gramfort
14h35 -- 14h55 : *The linear video coding paradigm for next generation video transmission*, M. Cagnazzo
14h55 -- 15h15 : *Multichannel audio source separation: variational inference of time-frequency sources from time-domain observations*, S. Leglaive

15h15 -- 15h45 : Café

15h15 -- 16h30 : Session Poster
16h30 -- 17h30 : *Débat: Recrutements à IDS: statut, priorités, rôle du LTCI*, G.Richard, I. Bloch

Exposés

François Portier (S2A)
A guided tour in Monte Carlo
Abstract

We review some popular Monte Carlo methods based on theoretical results and applications. We first present different variance reduction methods including Importance sampling and Control variates in the case of independent sampling. Second we discuss some extensions involving Markovian sampling.

Pietro Gori (IMAGES)

Parsimonious representations for modeling, visualizing and statistically analyzing brain tractograms

Abstract

Tractography from diffusion-weighted magnetic resonance imaging (DW-MRI) is the only non-invasive technique able to trace in vivo the wiring architecture of the human brain white matter. It results in a bundle of 3D polylines, usually called fibers, which are estimates of the trajectories of large groups of neural tracts. Recent methods produce up to one million of fibers, which makes it difficult to store, visualize and process them. In this talk, I will present an approximation scheme for tractograms which produces a parsimonious representation of weighted prototypes. The resulting approximation uses on average only 2% of the original fibers thus drastically reducing the computational burden for geometrical processes such as registration and atlas construction. In the last part of the talk, I will introduce a new project, called MetaTracts, which was selected as the Emergence DigiCosme project of this year.

Olivier Fercoq (S2A)

Restarting accelerated gradient methods with a rough strong convexity estimate

Abstract

Our main contribution is to show that, under a local quadratic growth condition, the restarted accelerated gradient and coordinate descent methods have a geometric rate of convergence for any restarting frequency. This allows us to take profit of restarting even when we do not know the strong convexity coefficient. The scheme can be combined with adaptive restarting to estimate on the fly the quadratic growth constant of the objective. Finally, we illustrate the properties of the algorithm on a regularized logistic regression problem and on a Lasso problem.

Jean-Marc Thierry (IMAGES)

User-centered modeling of skeletal weights for efficient 3D character animation

Abstract

This talk is about how to attach a 3D surface to a manipulation skeleton efficiently, through the definition of so-called skinning weights, which describe the amount of influence of the bone deformers over the vertices of the 3D mesh. We present a fast, robust and high-quality technique to skin a mesh with reference to a skeleton. We consider the space of possible skeleton deformations (based on skeletal constraints, or skeletal animations), and compute skinning weights based on an optimization scheme to obtain as-rigid-as-possible (ARAP) corresponding mesh deformations. We support stretchable-and-twistable bones (STBs) and spines by generalizing the ARAP deformations to stretchable deformers. Additionally, our approach can optimize joint placements.

If wanted, a user can guide and interact with the results, which is facilitated by an interactive feedback, reached via an efficient sparsification scheme. We demonstrate our technique on challenging inputs (STBs and spines, triangle and tetrahedral meshes featuring missing elements, boundaries, self-intersections, or wire edges).

Atef Ben Youssef (S2A)

Study of User Engagement in Spontaneous Human-Robot Interactions: Case of User Engagement Breakdown

Abstract

In this talk, I will present a study that investigates the use of multimodal data recorded while holding spontaneous interaction between humans and the humanoid robot Pepper.

The recorded dataset provides rich streams of data that could be used by research and development groups in a variety of areas.
Our study focus on user engagement breakdown defined as a failure to complete successfully the interaction and leaving before finishing it.
The goal is to detect as early as possible the user engagement breakdown during the interaction.

Henri Maitre (IMAGES)
L'esthétique dans les images
Abstract
Si la notion de qualité d'image est un sujet d'étude des traiteurs d'images depuis plus de quarante ans, les notions liées à l'esthétique des photographies ne sont apparues que depuis une dizaine d'années dans cette communauté. Mais avec l'automatisation des fonctions d'archivage et de recherche dans les bases de données numériques, les études qui sont consacrées à ce thème prolifèrent aujourd'hui, mettant à profit le développement des techniques d'apprentissage ainsi que les sites spécialisés dans l'archivage des photos.
Quel espoir peut-on mettre dans ces méthodes ? Quel profit tirent-elles de vingt siècle de littérature sur l'Esthétique, le Beau et l'Art tant en philosophie qu'en sociologie et en psychologie appliquée ? Comment se relient-elles aux travaux menés parallèlement en physiologie de la perception et en neurobiologie ?
On fera une brève présentation des diverses approches de l'esthétique des images en dégageant les critères qui peuvent se prêter à une analyse automatique. On examinera les deux grandes voies qui ont conduit à des algorithmes (l'approche objectiviste de Birkhoff et l'approche de "machine learning"), en insistant sur la seconde qui focalise aujourd'hui les attentions. On dressera un tableau critique des divers résultats aujourd'hui disponibles. On proposera quelques pistes pour le futur.

A. Gramfort (S2A)
Faster Independent Component Analysis (ICA) by preconditioning with Hessian approximation
Abstract
Independent Component Analysis (ICA) is a powerful technique for unsupervised data exploration that is widely used across fields such as neuroscience, astronomy, chemistry or biology. Linear ICA is a linear latent factor model, similar to sparse dictionary learning, that aims to discover statistically independent sources from multivariate observations. ICA is a probabilistic generative model for which inference is classically done by maximum likelihood estimation. Estimating sources by maximum likelihood leads to a smooth non-convex optimization problem where the unknown is a matrix called the separating or unmixing matrix. As the gradient of the likelihood is available in closed form, first order gradient methods, stochastic or non-stochastic, are often employed despite a slow convergence such as in the Infomax algorithm. While the Hessian is known analytically, the cost of its computation and inversion makes Newton method unpractical for a large number of sources. We show how sparse and positive approximations of the true Hessian can be obtained and used to precondition the L-BFGS algorithm. Results on simulations and two applied problems (EEG data and image patches) demonstrate that the proposed technique leads to convergence that can be orders of magnitude faster than algorithms commonly used today even when looking for hundred of sources.

M. Cagnazzo (MM)
The linear video coding paradigm for next generation video transmission
Abstract
Linear video coding (or pseudo-analog video coding) has recently attracted a lot of interest in the field of video transmission thanks to the very promising performance in the case of mobile multimedia distribution. Even though the basic idea is very simple (one removes all non-linear elements of a classical video transmission scheme and thus gets performance that gracefully follows the channel SNR), a thoroughful optimization involves some non-trivial problems (e.g. dimensionality reduction schemes, structured hermitian inverse eigenvalue problem, MIMO power allocation...). In this talk we give an overview of the framework and highlight some recent results.

Simon Leglaive (S2A)
Multichannel audio source separation: variational inference of time-frequency sources from time-domain observations
Abstract
A great number of methods for multichannel audio source separation are based on probabilistic approaches in which the sources are modeled as latent random variables in a Time-Frequency (TF) domain. For reverberant mixtures, it is common to approximate the time-domain convolutive mixing process as being instantaneous in the short-term Fourier transform domain, under a short mixing filters assumption. The TF latent sources are then inferred from the TF mixture observations. In this paper we propose to infer the TF latent sources from the time-domain observations. This approach allows us to exactly model the convolutive mixing process. The inference procedure relies on a variational expectation-maximization algorithm. We experimentally show that this approach allows us to improve the source separation quality in significant reverberation conditions.

Session Posters

Tom Dupre La Tour (S2A)
Non-linear Auto-Regressive Models for Cross-Frequency Coupling in Neural Time Series
Abstract
We address the issue of reliably detecting and quantifying cross-frequency coupling (CFC) in neural time series. Based on non-linear auto-regressive models, the proposed method provides a generative and parametric model of the time-varying spectral content of the signals. As this method models the entire spectrum simultaneously, it avoids the pitfalls related to incorrect filtering or the use of the Hilbert transform on wide-band signals. As the model is probabilistic, it also provides a score of the model ``goodness of fit'' via the likelihood, enabling easy and legitimate model selection and parameter comparison; this data-driven feature is unique to our model-based approach.
Using three datasets obtained with invasive electrophysiological recordings in humans and rodents, we demonstrate that these models are able to replicate previous results obtained with other metrics, but also reveal new insights such as the influence of the amplitude of the slow oscillation. Using simulations we demonstrate that our parametric method can reveal neural couplings with shorter signals than non-parametric methods. We also show how the likelihood can be used to find optimal filtering parameters, suggesting new properties on the spectrum of the driving signal, but also to estimate the optimal delay between the coupled signals, enabling a directionality estimation in the coupling.

Pavel Nikitin (work with J. Jung, M. Cagnazzo, B. Pesquet-Popescu) (MM)
Performance Study of View Synthesis with Small Baseline for Free Navigation
Abstract
In a typical Free Navigation service, view synthesis is expected to provide virtual views between the real captured views, in order to improve the smoothness of the navigation. Practical constraints prevent from capturing views with a very small baseline, so view synthesis is required. One way of synthesizing views is to use texture and depth information. It is of interest to understand how much current view synthesis technology is able to provide acceptable quality for synthesized views, in the framework of Free Navigation.
A new super multi-view content has been recently provided by the University of Brussels. This high-density content has the characteristic to have a very small baseline of 1mm and is particularly adapted for this study. In this study, some experiments of view synthesis with small baseline were performed. Experimental results are reported to understand how far view synthesis can be used, both

from an objective and from a subjective point of view. It was shown that according to subjective point of view, more views can be synthesized while maintaining acceptable quality. We suggest that objective (PSNR based) results can only be considered to rank 2 algorithms applied to the same configuration.

Anna Korba (S2A)

A learning theory of ranking aggregation

Abstract

Originally formulated in Social Choice theory, Ranking Aggregation, also referred to as Consensus Ranking, has motivated the development of numerous statistical models since the middle of the 20th century. Recently, the analysis of ranking/preference data has been the subject of a renewed interest in machine-learning, boosted by modern applications such as meta-search engines, giving rise to the design of various scalable algorithmic approaches for approximately computing ranking medians, viewed as solutions of a discrete (generally NP-hard) minimization problem. This paper develops a statistical learning theory for ranking aggregation in a general probabilistic setting (avoiding any rigid ranking model assumptions), assessing the generalization ability of empirical ranking medians. Universal rate bounds are established and the situations where convergence occurs at an exponential rate are fully characterized. Minimax lower bounds are also proved, showing that the rate bounds we obtain are optimal.

Theo Karagkioules (with Sami Mekki and Stefan Valentin) (MM)

HTTP adaptive streaming with indoors-outdoors detection in mobile networks

Abstract

In mobile networks, users may lose coverage when entering a building due to the high signal attenuation at windows and walls. Under such conditions, services with minimum bit-rate requirements, such as video streaming, often show poor Quality of Experience (QoE). We will present a Bayesian detector that combines measurements from two Smartphone sensors to decide if a user is inside a building or not. Based on this coverage classification, we will propose an HTTP adaptive streaming (HAS) algorithm to increase playback stability at a high average bitrate. Measurements in a typical office building show high accuracy for the presented detector and superior QoE for the proposed HAS algorithm.

Vedad Hulusic (MM)

Quality of Experience of Emerging Video Formats

Abstract

A key factor to determine the quality of experience (QoE) of a video is its capability to convey the large spectrum of perceptual phenomena that our eyes can sense in real life. In order to meet this demand, new video features, such as High dynamic Range (HDR) and Wide Color Gamut (WGC), High Frame Rate (HFR) and higher spatial resolutions (4K/8K), are employed. In addition, novel viewing paradigms, such as the 360-degree video, enable to make the visual experience more immersive. Our aim is to investigate the contribution of each of these technologies to QoE individually, but also to study their interaction, in order to quantify the benefits to users from their individual or joint effect. To this end, we conducted a series of studies, involving both subjective tests and objective measurements. Since Immersive video entails dealing with a huge amount of data in order to provide a credible sense of immersion, we study QoE in the realistic scenario where video is efficiently compressed in order to be distributed in the current and future networks.

Simon Leglaive (S2A)

Multichannel audio source separation: variational inference of time-frequency sources from time-domain observations

Abstract

A great number of methods for multichannel audio source separation are based on probabilistic approaches in which the sources are modeled as latent random variables in a Time-Frequency (TF)

domain. For reverberant mixtures, it is common to approximate the time-domain convolutive mixing process as being instantaneous in the short-term Fourier transform domain, under a short mixing filters assumption. The TF latent sources are then inferred from the TF mixture observations. In this paper we propose to infer the TF latent sources from the time-domain observations. This approach allows us to exactly model the convolutive mixing process. The inference procedure relies on a variational expectation-maximization algorithm. We experimentally show that this approach allows us to improve the source separation quality in significant reverberation conditions.

Pierre Ablin (S2A)

Beyond stochastic gradient for maximum likelihood based ICA on EEG

Abstract

Independent Component Analysis (ICA) is the process of finding a linear transform of signals which maximizes independence. This method is widely used in neuroscience for processing biological signals. We introduce a new algorithm for solving maximum likelihood Independent Component Analysis based on L-BFGS preconditioned with Hessian approximations. Using electroencephalographic data, we show that it can be orders of magnitude faster and that it solves the problem more accurately than the most widespread algorithm, Infomax. We demonstrate that that accuracy increase makes the biological conclusions drawn from the independent components more consistent and robust

Isabelle Bloch et al (IMAGES)

AI and IMAGES

Abstract

The aim of this presentation is to give an overview of the research activities of the IMAGES team in the field of artificial intelligence, and more specifically in machine learning, pattern, object and shape recognition, interaction, knowledge representation, decision and uncertainty management, and reasoning.

Sylvain Lobry (IMAGES)

Modèles markovien pour le traitement d'images satellitaires SAR

Abstract

This poster presents two water detection methods developed in the framework of the SWOT mission. SWOT is a future satellite mission which will measure the elevation of water surfaces using SAR interferometry. One of the key step towards this achievement is the detection of water bodies in the image. The particularities of the sensor used in the mission (Ka band with a near-nadir incidence angle) call for the development of dedicated methods.

In this poster, we present a method for the detection of large water bodies (i.e. lakes) and a second one for the detection of thin water bodies (i.e. rivers). The first one uses two Markov Random Fields (MRF) in an iterative process: one allows to obtain variable class parameters, the other uses these parameters to make the classification. The method dedicated to rivers detection combines a segment detection at a pixel level and an object-based method completing the detected segments (using Dijkstra's path finder and a MRF defined on the objects for the regularization).

Preliminary results are presented, showing a clear improvement compared to previous methods considered in the scope of the mission.

Adrien Kaiser (IMAGES)

Proxy Clouds for RGB-D Stream Processing

Abstract

Modern consumer depth cameras are widely used for 3D capture in indoor environments, for applications such as modeling, robotics or gaming. Nevertheless, their use is limited by their low resolution, with frames often corrupted with noise, missing data and temporal inconsistencies. In order

to cope with all these issues, we present Proxy Clouds, a multiplanar superstructure for real-time processing of RGB-D data. By generating a single set of planar proxies from raw RGB-D data and updating it through time, several processing primitives can be applied to improve the quality of the RGB-D stream or lighten further operations. We illustrate the use of Proxy Clouds on several applications, including noise and temporal flickering removal, hole filling, resampling, color processing and compression. We present experiments performed with our framework in indoor scenes of different natures captured with a consumer depth sensor.