Activity Report 2015

Project-Team GEOSTAT

Geometry and Statistics in acquisition data
# Table of contents

1. Members ................................................................. 1

2. Overall Objectives ..................................................... 2

3. Research Program ...................................................... 4
   3.1. Multiscale description in terms of multiplicative cascade 4
   3.2. Excitable systems and heartbeat signal analysis 7
   3.3. Speech analysis 8

4. Application Domains .................................................. 9

5. Highlights of the Year ................................................ 9

6. New Software and Platforms ......................................... 9
   6.1. Flux 9
   6.2. FluidExponents 9

7. New Results .......................................................... 9
   7.1. Super-resolution, multiscale data fusion and complex dynamics in Earth Observation and Universe Sciences 9
   7.2. Characterization of underlying stochastic dynamic of the cardiac muscle under fibrillation: singularity analysis and modeling 10
   7.3. Classification of Cardiac Arrhythmia in vitro based on Multivariate Complexity Analysis 10
   7.4. Classification of Cardiac Arrhythmia in vitro based on Multivariate Complexity Analysis 11
   7.5. Nonlinear trend removal should be carefully performed in heart rate variability analysis 11
   7.6. Quantification of Heart’s Recover by Multiscale Complexity Analysis of Heart Rate: a Validation Study 12
   7.7. Quantification and Its Approximate Solution of Action Potential in Neuron Models by Anharmonicity Analysis 15
   7.8. Image Reconstruction from Highly Corrupted Gradients 15
   7.9. Fast Image Edge-Aware Processing 16
   7.10. Low-Rankness Transfer for Realistic Denoising 16
   7.11. Turbulent Flow Estimation 17
   7.12. Pathological voice classification 17
   7.13. Emotion detection: project with Batvoice start-up 17

8. Bilateral Contracts and Grants with Industry .......................... 17

9. Partnerships and Cooperations ........................................ 17
   9.1. Regional Initiatives 17
   9.2. National Initiatives 18
   9.3. International Initiatives 18
      9.3.1.1. OPTIC 18
      9.3.1.2. Summary of work done in 2015 19

10. Dissemination ......................................................... 21
    10.1. Promoting Scientific Activities 21
        10.1.1. Scientific events organisation 21
        10.1.1.1. General chair, scientific chair 21
        10.1.1.2. Member of the conference program committees 21
        10.1.2. Journal 21
        10.1.3. Invited talks 21
        10.1.4. Scientific expertise 22
        10.2. Teaching - Supervision - Juries 22
          10.2.1. Teaching 22
          10.2.2. Supervision 22
          10.2.3. Juries 22
10.3. Popularization

11. Bibliography
Project-Team GEOSTAT

Creation of the Team: 2009 November 01, updated into Project-Team: 2011 January 01

Keywords:

Computer Science and Digital Science:
3.4.2. - Unsupervised learning
3.4.7. - Kernel methods
3.4.8. - Deep learning
5.3. - Image processing and analysis
5.3.2. - Sparse modeling and image representation
5.3.3. - Pattern recognition
5.3.5. - Computational photography
5.7. - Audio modeling and processing
5.7.3. - Speech
5.7.4. - Analysis
5.9. - Signal processing
5.9.2. - Estimation, modeling
5.9.3. - Reconstruction, enhancement
5.9.5. - Sparsity-aware processing

Other Research Topics and Application Domains:
2. - Health
2.2. - Physiology and diseases
2.2.1. - Cardiovascular and respiratory diseases
2.2.6. - Neurodegenerative diseases
3. - Environment and planet
3.3. - Geosciences
3.3.2. - Water: sea & ocean, lake & river
3.3.4. - Atmosphere

1. Members

Research Scientists
Hussein Yahia [Team leader, Inria, Researcher, HdR]
Khalid Daoudi [Inria, Researcher]
Oriol Pont [Starting Research position, until September 2015]

PhD Students
Camila Artana [Univ. Paris VI, from Oct 2015]
Hicham Badri [Inria]
Anass El Aouni [Rabat University, from Nov 2015]
Ayoub Tamim [Rabat University]

Post-Doctoral Fellows
Nicolas Brodu [Inria]
Guillaume Attuel [Inria]
2. Overall Objectives

2.1. Overall objectives

Singularity exponent A measure of the unpredictability around a point in a complex signal. Based on local reconstruction around a point, singularity exponents can be evaluated in different ways and in different contexts (e.g. non-localized, through the consideration of moments and structure functions, through the computation of singularity spectra). In GEOSTAT we study approaches corresponding to far from equilibrium hypothesis (e.g. microcanonical) leading to geometrically localized singularity exponents.

Singularity spectrum The mapping from scaling exponents to Hausdorff dimensions. The singularity spectrum quantifies the degree of nonlinearity in a signal or process, and is used to characterize globally the complexity of a signal.

Most Singular Manifold The set of most unpredictable points in a signal, identified to the set of strongest transitions as defined by the singularity exponents. From that set the whole signal can be reconstructed.

Adaptive Optics (AO) This term refers to a set of methodologies used, notably in Astronomical observations, to compensate for the loss of spatial resolution in optical instruments caused by atmospheric turbulence.

Fully Developed Turbulence (FDT) Turbulence at very high Reynolds numbers; systems in FDT are beyond deterministic chaos, and symmetries are restored in a statistical sense only.

Compact Representation Reduced representation of a complex signal (dimensionality reduction) from which the whole signal can be reconstructed. The reduced representation can correspond to points randomly chosen, such as in Compressive Sensing, or to geometric localization related to statistical information content (framework of reconstructible systems).

MMF Microcanonical Multiscale Formalism.

Sparse representation The representation of a signal as a linear combination of elements taken in a dictionary (frame or basis), with the aim of finding as less as possible non-zero coefficients for a large class of signals.

Universality class In theoretical physics, the observation of the coincidence of the critical exponents (behaviour near a second order phase transition) in different phenomena and systems is called universality. Universality is explained by the theory of the renormalization group, allowing for determination of the changes a physical system undergoes under different distance scales. As a consequence, different macroscopic phenomena displaying a multiscale structure (and their acquisition in the form of complex signals) can be grouped into different sets of universality classes.
GEOSTAT is a research project in **nonlinear signal processing**, with the fundamental distinction that it considers the signals as the realizations of complex dynamic systems. The driving approach is to understand the relations between complexity (or information content) and the geometric organization of information in a signal. For instance, for signals which are acquisitions of turbulent fluids, the organization of information is related to the effective presence of a multiscale hierarchy, of multifractal nature, which is strongly related to intermittency and multiplicative cascade phenomena; the determination of this geometric organization unlocks key nonlinear parameters and features associated to these signals; it helps understanding their dynamical properties and, as a consequence, their analysis. We use this approach to derive novel solution methods for **super-resolution** and data fusion in Universe Sciences acquisitions [10]. Another example can be found heartbeat signal analysis, where singularity exponents help understand the distribution of activation points in a signal during episodes of atrial fibrillation. Specific advances are obtained in GEOSTAT in using this type of statistical/geometric approach to get validated dynamical information of signals acquired in Universe Sciences, e.g. Oceanography or Astronomy. The research in GEOSTAT encompasses nonlinear signal processing and the study of emergence in complex systems, with a strong emphasis on geometric approaches to complexity. Consequently, research in GEOSTAT is oriented towards the determination, in real signals, of quantities or phenomena, usually unattainable through linear methods, that are known to play an important role both in the evolution of dynamical systems whose acquisitions are the signals under study, and in the compact representations of the signals themselves. Research in GEOSTAT is structured in two parts:

- Theoretical and methodological aspects in nonlinear signal processing.
- Applicative aspects which encompass biomedical data (heartbeat signal analysis with IHU LIRYC, biomedical applications in speech signal analysis) and the study of universe science datasets.

The theoretical objectives are:

- multiscale description in terms of multiplicative cascade (essential in the characterization of turbulent systems).
- Excitable systems (cardiac electrophysiology): study of intermittency phenomena.

The methodological tools used in reaching these objectives place GEOSTAT at the forefront of nonlinear signal processing and complex systems. We cite: singularity exponents [52][7][11], how these exponents can be related to **sparse representations** with reconstruction formulae [13] [53][13][13][5] and super-resolution in Oceanography and Earth Observation[10], comparison with embedding techniques, such as the one provided by the classical theorem of Takens [50], [42], the use of Lyapunov exponents [38], how they are related to intermittency, large deviations and singularity exponents, various forms of entropies, persistence along the scales, **optimal wavelets** [6], comparison with other approaches such as sparse representations and compressive sensing[14], and, above all, **the ways that lead to effective numerical and high precision determination of nonlinear characteristics in real signals**. Presently GEOSTAT explores new methods for analyzing and understanding complex signals in different applicative domains through the theoretical advances of the MMF, and the framework of **reconstructible systems** [51]. Derived from ideas in Statistical Physics, the methods developed in GEOSTAT provide new ways to relate and evaluate quantitatively the **local irregularity** in complex signals and systems, the statistical concepts of information content and most informative subset. That latter notion is developed through the notion of transition front and Most Singular Manifold. As a result, GEOSTAT is aimed at providing **radically new approaches** to the study of signals acquired from different complex systems (their analysis, their classification, the study of their dynamical properties etc.). A common characteristic of these signals, which is related to universality classes [45] [46] [43], being the existence of a multiscale organization of the systems. For instance, the classical notion of edge or border, which is of multiscale nature, and whose importance is well known in Computer Vision and Image Processing, receives profound and rigorous new definitions, in relation with the more physical notion of transition and fits adequately to the case of chaotic data. The description is analogous to the modeling of states far from equilibrium, that is to say, there is no stationarity assumption. From this formalism we derive methods able to determine geometrically the most informative part in a signal, which also defines its global properties and allows for **compact representation** in the wake of known problematic addressed, for instance, in **time-frequency analysis**. In this way, the MMF allows the reconstruction, at any prescribed quality threshold, of a
signal from its most informative (i.e. most unpredictable) subset, and is able to quantitatively evaluate key features in complex signals (unavailable with classical methods in Image or Signal Processing). It appears that the notion of transition front in a signal is much more complex than previously expected and, most importantly, related to multiscale notions encountered in the study of nonlinearity [48]. For instance, we give new insights to the computation of dynamical properties in complex signals, in particular in signals for which the classical tools for analyzing dynamics give poor results (such as, for example, correlation methods or optical flow for determining motion in turbulent datasets).

3. Research Program

3.1. Multiscale description in terms of multiplicative cascade

GEOSTAT is studying complex signals under the point of view of nonlinear methods developed statistical physics i.e. the methodologies developed to study complex systems, with a strong emphasis on multiresolution analysis. Linear methods in signal processing refer to the standard point of view under which operators are expressed by simple convolutions with impulse responses. Linear methods in signal processing are widely used, from least-square deconvolution methods in adaptive optics to source-filter models in speech processing. Because of the absence of localization of the Fourier transform, linear methods are not successful to unlock the multiscale structures and cascading properties of variables which are of primary importance as stated by the physics of the phenomena. This is the reason why new approaches, such as DFA (Detrended Fluctuation Analysis), Time-frequency analysis, variations on curvelets [49] etc. have appeared during the last decades. Recent advances in dimensionality reduction, and notably in Compressive Sensing, go beyond the Nyquist rate in sampling theory using nonlinear reconstruction, but data reduction occur at random places, independently of geometric localization of information content, which can be very useful for acquisition purposes, but of lower impact in signal analysis. One important result obtained in GEOSTAT is the effective use of multiresolution analysis associated to optimal inference along the scales of a complex system. The multiresolution analysis is performed on dimensionless quantities given by the singularity exponents which encode properly the geometrical structures associated to multiscale organization. This is applied successfully in the derivation of high resolution ocean dynamics, or the high resolution mapping of gaseous exchanges between the ocean and the atmosphere; the latter is of primary importance for a quantitative evaluation of global warming. Understanding the dynamics of complex systems is recognized as a new discipline, which makes use of theoretical and methodological foundations coming from nonlinear physics, the study of dynamical systems and many aspects of computer science. One of the challenges is related to the question of emergence in complex systems: large-scale effects measurable macroscopically from a system made of huge numbers of interactive agents [40], [37], [54], [44]. Some quantities related to nonlinearity, such as Lyapunov exponents, Kolmogorov-Sinai entropy etc. can be computed at least in the phase space [38]. Consequently, knowledge from acquisitions of complex systems (which include complex signals) could be obtained from information about the phase space. A result from F. Takens [50] about strange attractors in turbulence has motivated the determination of discrete dynamical systems associated to time series [42], and consequently the theoretical determination of nonlinear characteristics associated to complex acquisitions. Emergence phenomena can also be traced inside complex signals themselves, by trying to localize information content geometrically. Fundamentally, in the nonlinear analysis of complex signals there are broadly two approaches: characterization by attractors (embedding and bifurcation) and time-frequency, multiscale/multiresolution approaches. Time-frequency analysis [39] and multiscale/multiresolution are the subjects of intense research and are profoundly reshaping the analysis of complex signals by nonlinear approaches [36], [41]. In real situations, the phase space associated to the acquisition of a complex phenomenon is unknown. It is however possible to relate, inside the signal’s domain, local predictability to local reconstruction and deduce from that singularity exponents (SEs) [11] [7]. The SEs are defined at any point in the signal’s domain, they relate, but are different, to other kinds of exponents used in the nonlinear analysis of complex signals. We are working on their relation with:

- properties in universality classes,
• the geometric localization of multiscale properties in complex signals,
• cascading characteristics of physical variables,
• optimal wavelets and inference in multiresolution analysis.

The alternative approach taken in GEOSTAT is microscopical, or geometrical: the multiscale structures which have their "fingerprint" in complex signals are being isolated in a single realization of the complex system, i.e. using the data of the signal itself, as opposed to the consideration of grand ensembles or a wide set of realizations. This is much harder than the ergodic approaches, but it is possible because a reconstruction formula such as the one derived in [51] is local and reconstruction in the signal’s domain is related to predictability. This approach is analogous to the consideration of “microcanonical ensembles” in statistical mechanics.

Nonlinear signal processing is making use of various quantities related to predictability (for instance first Lyapunov exponent is related, from Oseledeč’s theorem, to the limiting behaviour of the response to perturbation in phase space [38]). In GEOSTAT our aim is to relate such classical quantities (among others) to the behaviour of SEs, which are defined by a limiting behaviour

\[ \mu(B_r(x)) = \alpha(x) r^{d+h(x)} + o(r^{d+h(x)}) \quad (r \to 0) \]

\[ d: \text{dimension of the signal’s domain}, \mu: \text{multiscale measure, typically whose density is the gradient’s norm}, \]

\[ B_r(x): \text{ball of radius } r \text{ centered at } x \). For precise computation, SEs can be smoothly interpolated by projecting wavelets:

\[ T_\Psi \mu(x, r) = \int_{\mathbb{R}^d} d\mu(x') \frac{1}{r^d} \Psi \left( \frac{x-x'}{r} \right) \]

\[ (\Psi: \text{mother wavelet, admissible or not}), \text{but the best numerical method in computing singularity exponents lies in the definition of a measure related to predictability [5]:} \]

\[ h(x) = \frac{\log T_{\Psi \mu}(x, r_0)/\langle T_{\Psi \mu}(\cdot, r_0) \rangle}{\log r_0} + o\left( \frac{1}{\log r_0} \right) \]

\[ \text{with: } r_0 \text{ is a scale chosen to diminish the amplitude of the correction term, and } \langle T_{\Psi \mu}(\cdot, r_0) \rangle \text{ is the average value of the wavelet projection (mother wavelet } \Psi) \text{ over the whole signal. Singularity exponents computed with this formula generalize the elementary "gradient’s norm" in a very statistically coherent way across the scales.} \]

SEs are related to the framework of reconstructible systems, and consequently to predictability. They unlock the geometric localization of a multiscale structure in a complex signal:

\[ \mathcal{F}_h = \{ x \in \Omega \mid h(x) = h \}, \]

\[ (\Omega: \text{signal’s domain}). This multiscale structure is a fundamental feature of a complex system, it is related to the cascade description of turbulence. We make use of this kind of description when analyzing turbulent signals: intermittency is observed within the inertial range and is related to the fact that, in the case of FDT, symmetry is restored only in a statistical sense, a fact that has consequences on the quality of any nonlinear signal representation by frames or dictionaries.
The example of FDT as a standard "template" for developing general methods that apply to a vast class of complex systems and signals is of fundamental interest because, in FDT, the existence of a multiscale hierarchy (i.e. the collection of sets $\mathcal{F}_h$ of equation 4) which is of multifractal nature and geometrically localized can be derived from physical considerations. This geometric hierarchy of sets is responsible for the shape of the computed singularity spectra, which in turn is related to the statistical organization of information content in a signal. It explains scale invariance, a characteristic feature of complex signals. The analogy from statistical physics comes from the fact that singularity exponents are direct generalizations of critical exponents which explain the macroscopic properties of a system around critical points, and the quantitative characterization of universality classes, which allow the definition of methods and algorithms that apply to general complex signals and systems, and not only turbulent signals: signals which belong to a same universality class share common statistical organization. In GEOSTAT, the approach to singularity exponents is done within a microcanonical setting, which can interestingly be compared with other approaches such that wavelet leaders, WTMM or DFA. During the past decades, classical approaches (here called "canonical" because they use the analogy taken from the consideration of "canonical ensembles" in statistical mechanics) permitted the development of a well-established analogy taken from thermodynamics in the analysis of complex signals: if $\mathcal{F}$ is the free energy, $T$ the temperature measured in energy units, $S$ the entropy and $\hat{\beta} = 1/T$, then the scaling exponents associated to moments of intensive variables $p \to \tau_p$ corresponds to $\beta \mathcal{F}$, $U(\hat{\beta})$ corresponds to the singularity exponents values, and $\delta(U)$ to the singularity spectrum. The singularity exponents belong to a universality class, independently of microscopic properties in the phase space of various complex systems, and beyond the particular case of turbulent data (where the existence of a multiscale hierarchy, of multifractal nature, can be inferred directly from physical considerations). They describe common multiscale statistical organizations in different complex systems [48], and this is why GEOSTAT is working on nonlinear signal processing tools that are applied to very different types of signals. For example we give some insight about the collaboration with LEGOS Dynbio team 1 about high-resolution ocean dynamics from microcanonical formulations in nonlinear complex signal analysis. Indeed, synoptic determination of ocean circulation using data acquired from space, with a coherent depiction of its turbulent characteristics remains a fundamental challenge in oceanography. This determination has the potential of revealing all aspects of the ocean dynamic variability on a wide range of spatio-temporal scales and will enhance our understanding of ocean-atmosphere exchanges at super resolution, as required in the present context of climate change. We show that the determination of a multiresolution analysis associated to the multiplicative cascades of a typical physical variable like the Sea Surface Temperature permits an optimal inference of oceanic motion field across the scales, resulting in a new method for deriving super resolution oceanic motion from lower resolution altimetry data; the resulting oceanic motion field is validated at super resolution with the use of Lagrangian buoy data available from the Global Drifter Program 2. In FDT, singularity exponents range in a bounded interval: $[h_\infty, h_{\max}]$ with $h_\infty < 0$ being the most singular exponent. Points $r$ for which $h(r) < 0$ localize the strongest transitions in the turbulent fluid, where an intensive physical variable like sea surface temperature behaves like $1/r^{h(r)}$. The links between the geometrically localized singularity exponents, the scaling exponents of structure functions, the multiplicative cascade and the multiscale hierarchy $\mathcal{F}_h$ is the following:

$$
\begin{align*}
\mathcal{F}_h &= \{ r \mid h(r) = h \} \\
D(h) &= \dim \mathcal{F}_h \\
\tau_p &= \inf_h \{ ph + 3 - D(h) \} \\
D(h) &= \inf_p \{ ph + 3 - \tau_p \}
\end{align*}
$$

Let $\mathcal{S}(x)$ be the bidimensional signal recording, for each sample point $x$ representing a pixel on the surface of the ocean of given resolution, the sea surface temperature (sst). To this signal we associate a measure $\mu$

---

whose density w.r.t Lebesgue measure is the signal’s gradient norm, and from which the singularity exponents are computed [6]. It is fundamental to notice here that, contrary to other types of exponents computed in Oceanography, such as Finite Size Lyapunov exponents, singularity exponents are computed at instantaneous time, and do not need time series.

Having computed the singularity exponents at each point of a SST signal, a microcanonical version of the multiplicative cascade associated to the scaling properties of the SST become available. The idea of the existence of a geometrically localized multiplicative cascade goes back to [47]. The multiplicative cascade, written pointwise, introduces random variables \( \eta_{l/l'}(x) \) for \( 0 < l' < l \) such that

\[
\eta_{l/l'}(x) = \eta_{l/l'}(x) \eta_{l/l'}(x)
\]

\[
\tau_{0}(x,l') = \eta_{l/l'}(x) \tau_{0}(x,l)
\]

in which the equality is valid pointwise and not only in distribution. Any mother wavelet \( \psi \) such that the process \( \eta_{l/l'}(x) \) is independant of \( \tau_{0}(x,l) \) is called an optimal wavelet: it optimizes inference of physical variables across the scales and consequently describes the multiplicative cascade at each point \( x \) in the signal domain. The injection variables \( \eta_{l/l'}(x) \) are indefinitely divisible: \( \eta_{l}(x) \eta_{l'}(x) = \eta_{l+l'}(x) \). It is possible to optimize cross-scale inference of physical variables by considering a multiresolution analysis associated to a discrete covering of the "space-frequency" domain. Denoting as usual \( (V_j)_{j \in \mathbb{Z}} \) and \( (W_j)_{j \in \mathbb{Z}} \) the discrete sequence of approximation and detail spaces associated to a given scaling function, and denoting by \( \psi \in L^2(\mathbb{R}^2) \) a wavelet which generates an Hilbertian basis on each detail space \( W_j \), it is known that the detail spaces encode borders and transition information, which is ideally described in the case of turbulent signals by the singularity exponents \( h(x) \). Consequently, a novel idea for super-resolution consists in computing a multiresolution analysis on the signal of singularity exponents \( h(x) \), and to consider that the detail information coming from spaces \( W_j \) is given the signal \( h(x) \). The associated orthogonal projection \( \pi_j : L^2(\mathbb{R}^2) \to W_j \) defined by \( \pi_j(h) = \sum_{n \in \mathbb{Z}} \langle h | \psi_{j,n} \rangle \psi_{j,n} \) is then used in the reconstruction formula for retrieving a physical variable at higher resolution from its low resolution counterpart. If \( \mathcal{G}(x) \) is such a variable, we use a reconstruction formula: \( A_{j-1} \mathcal{G} = A_j \mathcal{G} + \pi_j h \) with \( A_j : L^2(\mathbb{R}^2) \to V_j \) is the orthogonal projection on the space \( V_j \) (approximation operator) and \( \pi_j \) is the orthogonal projection on the detail spaces \( W_j \) associated to the signal of singularity exponents \( h(x) \). Validation is performed using Lagrangian buoy data with very good results [10]. We have realized a demonstration movie showing the turbulent ocean dynamics at an SST resolution of 4 km computed from the SST microcanonical cascade and the low-resolution GEKCO product for the year 2006 over the southwestern part of the Indian Ocean. We replace the missing data in the SST MODIS product (clouds and satellite swath) by the corresponding data available from the Operational SST and Sea Ice Analysis (OSTIA) provided by the Group for High-Resolution SST Project [11], which, however, is of lower quality. Two images per day are generated for the whole year of 2006. The resulting images show the norm of the vector field in the background rendered using the line integral convolution algorithm. In the foreground, we show the resulting vector field in a linear gray-scale color map. See link to movie (size: 800 Mo).

### 3.2. Excitable systems and heartbeat signal analysis

In collaboration with IHU LIRYC, we are developing novel approaches to heartbeat signal analysis for understanding chronic atrial fibrillation. The noisy aspect of data recorded by electrodes, on the inner surface of human atria during episodes of atrial fibrillation, exhibit intriguing features for excitable media. Instead of phase chaos as typically expected, it shares many common traits of non-equilibrium fluctuations in disordered systems or strong turbulence. To assess those peculiar observations we investigate a synaptic plasticity that affects conduction properties. Electrical synapses comprise many different kinds of connexins, which may be affected by diverse factors, so we use a generic approach. Slight detuning of their linear response leads to an instability of the modulating agents, here an excess charge. Acting on slow time scales of repolarisation, it is understood as collective modes propagating through and retroacting on each synapse: the medium is desynchronised. It is not a syncytium. We propose to associate transient states with a phenomenon called...
**electrical remodelling**, which has not received any accepted description thus far. Moreover, from the properties of the model it is possible to start exploring phase space. Transitions between different regimes could help decipher stages in the evolution of the disease from acute to chronic, one main goal of cardiovascular research.

Theoretically, a myocardium is an excitable tissue acting under normal circumstances as a functional syncytium of myocardial cells. Models of excitability for the heart are reaction-diffusion systems describing the propagation of electric pulses called action potentials similarly to models for axons. Reaction results from ionic exchange cycles between the cytoplasm of excitable cells and their extra-cellular medium, when initiated by a stimulus above some threshold. Pulses are robust topological structures.

Considering the stable fixed point as a phase resetting state, even under periodic forcing, no return map can be drawn, therefore no chaos. Chaos may arise in temporal sequences. As propagation becomes further involved in the arrhythmia, spatio-temporal chaos amounts to the breakup of spirals from direct fore front and back front collisions within the pulse train. This is the paradigm for cardiac fibrillation.

The distributions of amplitudes all collapse on a scaling function $G$. We map exponents on data patients provided by IHU LIRYC showing non-universal properties. Singular exponents are observed with consistent Hausdorff dimension of sets $D(h)$. Negative contribution is high, suggesting an underlying multiplicative process.

Excess charge in cells like of $Ca$ may perturb the dynamics of synapses. We consider a physiologically plausible linear response of synapses to the electro-chemical potential. This response is unknown as of today. The new dynamics may interact with excitability. It has the specific form of a Rayleigh instability. Cycles become retarded or advanced. Hopf bifurcation and chaos are allowed creating EADs (Early After Depolarization). Regarding propagation, pulses are pinned and released on a chaotic background. Cycle modulations create defects via facilitation through the third dimension. Defects proliferate creating a glassy phase, which back-scatter fronts in 1D and roughens them in 2D. Further effective inhibitor diffusion splits them. Electrical remodelling is the time by which the background has pervaded the tissue.

There are features of Self Organized Criticality (SOC) in large regions of phase space. Pulses have a phase and propagate on a random medium. For instance one paradigm we investigate would be:

$$\partial_t \theta + \sin(\theta + \tilde{\phi}) = \Omega + \partial_{xx} \theta$$

$(\theta$: phase of activation front, $\Omega$: tachycardia frequency, $\tilde{\phi}$: phase perturbation). Randomness realtalisces non-linearly, which tells that the noise is quenched and reset. For instance in $1 + 1D$, spatio-temporal maps look very much like optimal directed paths along diagonals. In $1 + 2D$, we are guessing that pulses do propagate in the (q)KPZ universality class, just as the remodelling front does. We will need refined numerical computations to confirm this. This is consistent with a multiplicative process. Physiologically, one interesting bonus is the interpretation of non-reentrant Tachycardia as dislocation patterns slowly evolving.

### 3.3. Speech analysis

Our research in speech processing focus on the development of novel nonlinear analysis methods for the characterization and classification of pathological and affective speech. For the latter, classical linear methods do not generally capture the nonlinearity, aperiodicity, turbulence and noise that can be present in pathological voices. We thus aim to design and extract new features that allow better characterization/classification of such voices, while being easy to interpret by clinicians. For the former, recent research have shown that the voice source signal information allow significant improvement of speech emotion detection systems. Our goal is to develop novel nonlinear techniques to extract relevant voice source features and to design efficient machine learning algorithms for robust emotion classification.
4. Application Domains

4.1. Application Domains

Application aspects in GEOSTAT encompass biomedical data (heartbeat signal analysis with IHU LIRYC, biomedical applications in speech signal analysis) and the study of universe science datasets. GEOSTAT's objectives in analysis of biomedical data hinge on the following observations:

• The analysis and detection of cardiac arrhythmia and pathological voice disorders is a paradigm in nonlinear methodologies applied to these types of signals.
• The classical hypothesis under linear approaches are confronted with strong nonlinearities, aperiodicity and chaotic phenomena present in these signals.
• Existing nonlinear approaches are lacking physiological interpretation. Our objective in this part is to propose new measures based on low-level transition characteristics, these transition phenomena being related to general concepts associated to predictability in complex systems.

5. Highlights of the Year

5.1. Highlights of the Year

• Article published on Inria web site: link to page on Nicolas Brodu’s Nature Communications paper: Spanning the scales of granular materials through microscopic force imaging, [17].

6. New Software and Platforms

6.1. Fluex

**KEYWORDS:** Signal - Signal processing

**SCIENTIFIC DESCRIPTION**
Fluex is a nonlinear signal processing software for 1D, 2D 3D and 3D+t general signals.

**FUNCTIONAL DESCRIPTION**

• Fluex is a library in nonlinear signal processing, written in C++, developed under Gforge, able to analyze turbulent and natural complex signals.
• Fluex is able to determine low level features in natural signals that cannot be determined using standard linear techniques.
• Participants: Hussein Yahia, Denis Arrivault, Rémi Paties.
• Contact: Hussein Yahia.
• URL: https://geostat.bordeaux.inria.fr/index.php/downloads.html.
• Fluex is deposited APP, Inter Deposit Digital Number: IDDN.FR.001.51.0028.000.S.P.2015.000.21000

6.2. FluidExponents

• Participants: Hussein Yahia and Antonio Turiel
• Contact: Hussein Yahia
• URL: https://geostat.bordeaux.inria.fr/index.php/downloads.html.

7. New Results

7.1. Super-resolution, multiscale data fusion and complex dynamics in Earth Observation and Universe Sciences

**Participants:** Hussein Yahia, Nicolas Brodu, Guillaume Attuel, Sylvain Bontemps, Nicola Schneider, Camila Artana, Dharmendra Singh, Joel Sudre, Véronique Garçon, Christine Provost, Anass El Aouni, Oriol Pont, Khalid Daoudi, Ayoub Tamim, Akankhsa Garg, Frédéric Frappart, Luc Bourrel.
In these thematics the following research is started or continued:

- Super-resolution and data fusion in Earth Observation. Important results obtained in validation either in ocean dynamics or partial $\rho CO_2$ pressures in ocean/atmosphere exchanges, coastal upwelling.
- Development of a new super-resolution model for multispectral images, demonstration on MODIS (NASA) and Sentinel-2 (ESA) data.
- Adaptive optics.
- Starting of a strong collaboration with Laboratoire d’Astrophysique de Bordeaux on the dynamics of galactic clouds.
- Supervised classification of ground terrain through multispectral imagery (with OPTIC associated team).
- Anomaly detection in SAR images, application to flood monitoring in Equator.
- Starting of a project on dune monitoring.

Publications: [21], [19], [29], [22], [25], [18] A. Tamim’s PhD HAL link, IEEE TGRS article on AO [21].

7.2. Characterization of underlying stochastic dynamic of the cardiac muscle under fibrillation: singularity analysis and modeling

**Participants:** Guillaume Attuel, Binbin Xu, Oriol Pont, Hussein Yahia.

Signals of heart electrical activity obtained through invasive measurements show properties not compatible with the purely excitable nature of cellular dynamics. We have developed a synaptic perturbation model of that dynamics showing good properties. Perturbations propagate an inter-cell desynchronization formally like diffusion-coupled chaotic maps. The model enters the universality class of directed propagative fronts of the type random branching or directed polymer in a disordered medium in $1+1D$ and pinning-depinning contact lines in $1+2D$. In the continuum limit, the universality class is supposed to be the one of KPZ (Kardar Parisi Zhang) or VM (Voter Model). This is a change in paradigm for the description of cardiac dynamics. We make use of this hypothesis to characterize precisely the state of the substrate through appropriate signal analysis, with the goal of being able to distinguish between different states or types in the pathology. We are involved in a technological transfer on this activity since summer 2015.

Publication: [26].

7.3. Classification of Cardiac Arrhythmia in vitro based on Multivariate Complexity Analysis

**Participant:** Binbin Xu.

**Background:** The animal models (in vitro or in vivo) provide an excellent tool to study heart diseases, among them the arrhythmia remains one of the most active research subjects. It can be induced or treated by drugs, electrical stimulation, hypothermia etc. **Problems:** However, the inducing or treating effects in cardiac culture often happen long after the initial applications or in some relatively short time windows. So, it is necessary to capture and classify rapidly the signal change. Human-assisted monitoring is time-consuming and less efficient. An automatic classification method for real-time use would be useful and necessary. **Methods:** Since electrocardiographic signals are features by repetitive or similar patterns reflecting the intrinsic information about the patient (or culture), analyzing these patterns could help not only to monitor the status’s change but also to evaluate/explore the physiologic control mechanisms. Methods based on complexity analysis are of considerable interest in this case. **Aims:** Compare different complexity analysis methods in order to find the most appropriate ones to discriminate the normal cardiac signals from arrhythmic ones acquired from a cardiac cell culture in vitro. The selected features are then used by a SVM classifier.
Results: Among the six complexity analysis methods, Time Lagging (T.Lag) method allowed obtaining the best discrimination index (normal vs. arrhythmic, \( p \)-value, \( 9 \times 10^{-23} \)). The proposed Modified Hurst Exponent (\( H_{\text{ExM}} \)) showed better performance than original Hurst Exponent with well-improved \( p \)-value (from \( 0.019 \) to \( 2 \times 10^{-9} \)). The Approximate Entropy (ApEn), Sample Entropy (SampEn) and Detrended Fluctuation Analysis gave good discrimination ratio but with larger \( p \)-values (at order \( 10^{-3} \)). Combination of T.Lag, \( H_{\text{ExM}} \) and ApEn can provide a more robust classifier and allow monitoring and classifying in an automatic way the electrical activities’ changes in the cardiac cultures.
Publication: [28].

7.4. Classification of Cardiac Arrhythmia in vitro based on Multivariate Complexity Analysis

Participant: Binbin Xu.

Physiological signals are temporal series containing a lot of information, and their analysis (either for diagnosis or evolution monitoring) necessitates tools that take into account their intrinsic characteristics, notably in terms of unpredictability and high number of parameters. Methodologies coming from chaotic and nonlinear dynamical systems contain some useful building blocks in that perspective, and allow a qualitative link with phenomenological and bio-inspired models. The objective of this work is to introduce some methods in nonlinear dynamics useful for the processing of these types of signals. An application of these tools is illustrated in the processing of potential electrical fields acquired from in vitro culture cells on newborn rats. Both normal (regular contraction of cells) and arrhythmic (disordered contractions) cases are contemplated.

The bifurcation diagram is an example of a tool that can be used in the temporal analysis of an experimental system.
Publication: [32].

7.5. Nonlinear trend removal should be carefully performed in heart rate variability analysis

Participants: Binbin Xu, Oriol Pont, Hussein Yahia, Rémi Dubois.
Figure 2. CPE period bifurcation diagram. Cells are stimulated by M1 electrode during 5 minutes with an impulsion train of 200 um and frequency 100 Hz. Three particular phenomena in cell behaviour: A (t = 1 hour) chaotic state, B (t = 3 hour) and C (t = 3.5 hour) period doubling phase, D (t = 4 hour) regular and stable rhythm.

Background: In Heart rate variability analysis, the rate-rate time series suffer often from aperiodic non-stationarity, presence of ectopic beats etc. It would be hard to extract helpful information from the original signals. Problem: Trend removal methods are commonly practiced to reduce the influence of the low frequency and aperiodic non-stationary in RR data. This can unfortunately affect the signal and make the analysis on detrended data less appropriate. Objective: Investigate the detrending effect (linear & nonlinear) in temporal / nonlinear analysis of heart rate variability of long-term RR data (in normal sinus rhythm, atrial fibrillation, congestive heart failure and ventricular premature arrhythmia conditions). Methods: Temporal method: standard measure SDNN; Nonlinear methods: multi-scale Fractal Dimension (FD), Detrended Fluctuation Analysis (DFA) & Sample Entropy (SampEn) analysis.

Results: The linear detrending affects little the global characteristics of the RR data, either in temporal analysis or in nonlinear complexity analysis. After linear detrending, the SDNNs are just slightly shifted and all distributions are well preserved. The cross-scale complexity remained almost the same as the ones for original RR data or correlated. Nonlinear detrending changed not only the SDNNs distribution, but also the order among different types of RR data. After this processing, the SDNN became indistinguishable between SDNN for normal sinus rhythm and ventricular premature beats. Different RR data has different complexity signature. Nonlinear detrending made the all RR data to be similar, in terms of complexity. It is thus impossible to distinguish them. The FD showed that nonlinearly detrended RR data has a dimension close to 2, the exponent from DFA is close to zero and SampEn is larger than 1.5 – these complexity values are very close to those for random signal. Conclusions: Pre-processing by linear detrending can be performed on RR data, which has little influence on the corresponding analysis. Nonlinear detrending could be harmful and it is not advisable to use this type of pre-processing. Exceptions do exist, but only combined with other appropriate techniques to avoid complete change of the signal’s intrinsic dynamics. One submitted publication.
Figure 3. Complexity Space based on FD, DFA & SampEn, for RR data in NSR, AF, CHF & VPB conditions.
7.6. Quantification of Heart’s Recover by Multiscale Complexity Analysis of Heart Rate: a Validation Study

Participants: Binbin Xu, Hussein Yahia, Rémi Dubois.

Background: Heart rate analysis is the common analysis of heart’s function. After the drug treatment of cardiac arrhythmia, the heart rate looks like the same as the group with normal sinus rhythm.

Figure 4. Heart rate RR time series in three cases: (1) arrhythmia, pre-treatment; (2) arrhythmia suppressed, post-treatment; (3) normal healthy group.

Problem: However, the visibly "same dynamics" for post-treatment & normal group does not reflect the true intrinsic dynamics of the heart. Methods: Using multi-scale complexity analysis to quantify and qualify the heart rate’s dynamics.

Figure 5. Analysis of heart rate variation by multiscale (coarse-graining) fractal dimension
Results: Thought the analysis shown in time domain that the dynamics of post-treatment and normal group looked similar. Their dynamics is completely different: (1) for normal heart rate, the multiscale fractal dimension is almost linearly decreased — invariance; (2) for arrhythmic heart rate before and after treatment, they converged to a certain value. All these suggested that item after the drug treatments, the heart’s function is not still fully restored and more recovering time is needed. The multiscale complexity analysis can be used to quantify the heart function’s recovering and optimize the post-treatment.

One submitted publication.

7.7. Quantification and Its Approximate Solution of Action Potential in Neuron Models by Anharmonicity Analysis

Participants: Binbin Xu, Hussein Yahia, Rémi Dubois.

Action potential (AP) plays an important role to initiate and maintain the cell-cell communication. The nerve impulses are extensively studied but the action potential is less investigated as in other types of cells (for example, cardiac action potential). The AP can tell more about the state of the cell. It reflects the physical / chemical intracellular exchanges. Any changes in the cell would change the form/geometry of AP, or a more relevant term the harmony. The intrinsic changes would modify the harmony of the impulses train. The broken harmony (form/geometry change) of the impulse train means that there would be some problems in the cell. This provides an indirect way to study the intrinsic dynamics the cells.

In the work of P. Hanusse proposed a very interesting signal analyzing approach by anharmonicity, especially for signals with nonlinear oscillations properties exhibited in many physical / biological systems. This is exactly the case for neuron impulses trains. The principle is to describing the signal with their harmonic behaviors by solving the nonlinear phase equation. The obtained phase is thus used to reconstruct a solution of the original signal. The key notion is the nonlinear trigonometry that they developed. According to this approach, for any periodic signal \( x(t) = x_0 + x_1 \cos(\phi(t)), \) its phase can be obtained by the proposed general solution \( t(\phi) = \phi + \sum_{k=1}^{n} a_k \text{hpsin}_1(\phi - p_k, r_k) - b_k \text{hpcos}_1(\phi - p_k, r_k) \) which can be used to reconstruct the original signal \( x(t). \)

There is no practical implementation in their papers. Here we propose a first order solution of the original analytical equation. It can be used to quantify the harmonicity of the action potential.

\[
\begin{align*}
\text{hpsin}_1 &= -i \left( \frac{- \ln (1 - e^{it}r) + \ln (1 - re^{-it})}{2r} \right), \\
\text{hpcos}_1 &= - \frac{\ln (1 - e^{it}r) + \ln (1 - re^{-it})}{2r}
\end{align*}
\]

The phase of a signal can be solved as \( \phi(t) = t_0 + a_1 \text{hpsin}_1(t - t_1, r) - a_2 \text{hpcos}_1(t - t_2, r), \) so the signal can be reconstructed as \( x(t) = \cos(\phi(t)). \) The related parameters can be obtained by regression or nonlinear optimization methods. In consequence, all AP can be quantified by the anharmonicity parameter \( r. \)

FitzHugh-Nagumo (FHN) model has been one of the basic models to study the action potential’s dynamics. It’s derived from the Hodgkin-Huxley model and is physiologically correct. We take here dynamics of APs from FHN model to illustrate the anharmonicity analysis. As shown in the following figure, anharmonicity analysis is far more efficient than Fourier representation. Even with 8 Fourier terms, the signal is ill-represented with Gibbs phenomena. In contrast, 3 anharmonic terms exhibit a quasi-identical fit. These AP are quantified with an anharmonicity \( r = 0.7384. \)

More development of anharmonicity analysis on AP is ongoing in order to provide a more efficient way to generalize the practices and for better solutions. We believe that anharmonicity analysis can help to quantify/qualify the AP in a different yet efficient way than conventional analysis.

One submitted publication.

7.8. Image Reconstruction from Highly Corrupted Gradients

Participants: Hicham Badri, Hussein Yahia, Driss Aboutajdine.
Surface-from-Gradients (SfG) is an important step in many imaging applications. It consists in reconstructing an image/surface from corrupted gradient fields, which results in an ill-posed problem. We propose to use sparsity to regularize the problem. The first approach uses sparsity in the gradient field together with a robust norm on the data-fitting term and was presented at CVPR 2014. The new approach uses a non-local regularization that manipulates non-local similar patches of the corrupted gradient and forcing them to be low-rank. The two approaches significantly outperform previous optimization-based SfG methods on both synthetic and real data.

One submitted publication.

7.9. Fast Image Edge-Aware Processing

Participants: Hicham Badri, Hussein Yahia, Driss Aboutajdine.

We present a framework for fast edge-aware processing of images and videos. This is an extension of our previous SIGGRAPH Asia 2013 paper. The proposed approach uses non-convex sparsity on the gradients of the latent smooth image to better preserve sharp edges. We develop tools based on first order proximal estimation for fast processing. We also propose fast and efficient numerical solutions based on separable filters estimation, which enables our method to perform fast high-quality smoothing on large-scale images. Extensive experiments show that the proposed method produces high-quality smoothing compared to state-of-the-art methods, while being fast and simple to implement.

Publication: [15].

7.10. Low-Rankness Transfer for Realistic Denoising

Participants: Hicham Badri, Hussein Yahia.

Image restoration is a very challenging task in low-level vision and is extensively used in many imaging applications. Sparsity in various forms (dictionary learning, low-rank estimation,...) has shown to be the key for successful image denoising. However, the standard noise model used to validate the results is mainly Gaussian and uniform, with known standard deviation. Unfortunately, these assumptions do not hold for real camera noise. Instead of using sparsity to model the singular values of non-local clean similar patches, we use a learning model that trains a mapping between the noisy and ground-truth clean singular values. The training is performed on real camera noise, contrary to previous methods. Experiments show that the proposed method significantly outperforms previous denoising works on real non-uniform noise and does not require estimating the standard deviation of the corruption.

One publication accepted with minor revision at IEEE Transactions on Image Processing, publication date: 2016.
7.11. Turbulent Flow Estimation

**Participants:** Hicham Badri, Hussein Yahia.

We use singularity exponents (SE) to regularize the problem of turbulent flow estimation under the assumption that the brightness constancy constraint holds also for (SE). We also use weighted filtering (Lucas–Kanade’s solution) and sparsity on the data-fitting term to improve robustness to outliers. The proposed motion estimation is built on a Gaussian pyramid and uses the theory of warping for a better estimation of large displacements. Experiments on synthetic data show that the proposed method outperforms sophisticated methods while being simple.

Work in progress.

7.12. Pathological voice classification

**Participants:** Khalid Daoudi, Nicolas Brodu.

Based on our GCI detection algorithm, we redefined the classical pitch perturbation measures that are widely used in voice quality assessment. We showed that our perturbation measures yield significantly better performance in pathological voice classification than classical measures. We also showed that some matching pursuit features can allow good performances in discrimination between pathological voice categories.

Publications: [31], [30].

7.13. Emotion detection: project with Batvoice start-up

**Participants:** Khalid Daoudi, Nicolas Brodu.

Geostat has been granted in 2015 a Carnot-Inria contact to fund a 1 year engineer to develop a prototype of a speech emotion detection system.


**Participants:** Hussein Yahia, Guillaume Attuel, Oriol Pont, Binbin Xu.

Geostat has been granted in 2015 a fund from Inria DGT to conduct pre-clinical validation from patient database acquired by IHU LIRYC.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

- Geostat has been granted in 2015 a Carnot-Inria contact to fund a 1 year engineer to develop a prototype of a speech emotion detection system. This contact, led by K. Daoudi, is in collaboration with the start-up BatVoice which targets the commercialization of affect-interactive digital systems.
- Geostat has set up an industrial contract with LECTRA Company, on the development of nonlinear signal processing tools for analysis signals acquired from turbines.
- DGT Inria has funded a Proof of Concept on heartbeat analysis with IHU LIRYC.
- Patent: Geostat is in the process of depositing a patent on heartbeat signal analysis with the help of Inria Transfer and the Cabinet Netter.

9. Partnerships and Cooperations

9.1. Regional Initiatives

Conseil Regional Aquitaine Project Cavernom (ref. 9129): Cardiac Arrhythmia Complexity and Variability by Means of Robust Nonlinear Methods. One year.
9.2. National Initiatives


- PhD grant provided by CNES and Conseil Regional Aquitaine, in collaboration with Laboratoire d’Astrophysique de Bordeaux. Starting: end 2016. Subject: understanding the dynamics of galactic dust clouds and their relation with star formation process.

- PhD grant for C. Artñana from UPMC University, under co-supervision with H. Yahia and C. Provost (LOCEAN, Paris).

9.3. International Initiatives

- The Toubkal project "Caractérisation multi-capteurs et suivi spatio-temporel de l’Upwelling sur la côte atlantique marocaine par imagerie satellitaire", led by K. Daoudi, has been accepted. The partners in this project are: Faculté des sciences de Rabat, Centre Royal de Télédetection Spatiale, Mercator-Ocean and Geostat.

9.3.1. Inria Associate Teams not involved in an Inria International Labs

9.3.1.1. OPTIC

Title: Optimal inference in Complex and Turbulent data.

International Partner (Institution - Laboratory - Researcher):

IITR (India), Department of Electronics and Communication Engineering: Dharmendra Singh

Start year: 2014

See also: https://optic.bordeaux.inria.fr/.

The OptIC associated team is co-managed by Prof. D. Singh (IIT Roorkee) and N. Brodu, H. Yahia (Inria Geostat).

The OptIC associated team targets the extension and development of a strong collaboration between Inria Geostat team and Indian Institute of Technology Roorkee, Dept of Electronics and Computer Engineering (Prof. D. Singh’s group) on non-linear Signal Processing for Universe Sciences, with a strong emphasis on data fusion in Earth Observation and monitoring. Non-linear Physics puts strong evidence of the fundamental role played by multiscale hierarchies in complex and turbulent data: in these data, the information content is statistically localized in geometrical arrangements in the signal’s domain, while such geometrical organization is not attainable by classical methods in linear signal processing. This is one of the major drawbacks in the classical analysis of complex and turbulent signals. The goal of this associated team is to show that inference of physical variables along the scales of complex and turbulent signals can be performed through optimal multiresolution analysis performed on non-linear features and data extracted from the signals, resulting in novel and powerful approaches for data fusion between different acquisitions (in temporal/spatial/spectral resolutions). This program needs both strong expertise in the physical processes beyond the acquisitions and the application of non-linear physics ideas on the behavior of the acquired physical phenomena. The proposal will focus on specific applications in Earth Observation and monitoring for which the Indian partner has developed a very strong expertise, notably in its knowledge and use of the physical processes in remote sensing acquisitions. This partnership is an extremely interesting and high potential collaboration between two teams which focus separately either on the acquisition of the physical processes or their analysis by Complex Systems and non-linear physics methodologies. The recent results obtained in super-resolution by GEOSTAT promises strong applications to a much wider range of Universe Sciences problems, notably with a strong emphasis on data fusion between the physical variables acquired on related
but different acquisitions. OptiC builds on a collaboration between Inria and IIT ROORKEE teams, added with partners in Universe Sciences and earth observation (ONERA, CNRS) already involved in research actions with Geostat.

9.3.1.2. Summary of work done in 2015

Figure 7. Left: nearest neighbors classification with the original 500m data, showing a cross-validated accuracy of $0.81 \pm 0.03$. Right: results on the superresolved 250m data, accuracy of $0.83 \pm 0.02$. Although the performance did not improve significantly on the reference points, the generalization capabilities are greatly enhanced: water regions (river, right, and canal, left) are well recognized, together with villages (magenta dots) and bare soil (red) adequately spotting the landscape instead of the incorrect zones on the lower-left at 500m. The other two classes are crops/small vegetation (yellow) and dense/tropical vegetation (green). The region is a MODIS sinusoidal projection around Roorkee.

Figure 8. Classified MODIS image (a) Low resolution (b) after resolution enhancement.

- The public availability of low-resolution MODIS data is cost-effective, but limited in precision. Some applications, such as land monitoring and anomaly detection, must not only operate on objects smaller than provided in the freely available data, but also offer a high level of confidence in the classified land occupation. We are working on both aspects:
- Augmenting numerically the resolution of the images. This can be done with different methods, out
of which two are currently explored in our team. The first relies on wavelet decompositions, with an attempt at preserving the spatial structure around each pixel (e.g. edges). This is typically done by propagating the high-frequency components to higher wavelet decomposition level through some interpolation mechanism, plus artifact-reduction steps. The second method considers a sub-pixel mixing model which is fit from all multispectral bands. The assumption is that, irrespectively of the reflectance of natural elements at each wavelength, the proportion of these elements is a physical property shared through all spectral bands. Then, unmixing is performed in order to estimate the best sub-pixels. See figure 7.

- Resolution-augmented images are then exploited for classification. We use by field measurements, in order to provide the ground truth for a corpus of well-registered locations, which together encompass a wide variety of objects (e.g. urban, crops, etc). We then train our super-resolution algorithms, and quantitatively assess our super-resolved maps on how well they improve the performance of land classification. As the final accuracy results from the interplay between the considered feature space and the classification method itself, we quantify both aspects with cross-validated data sets. We have tested state of art classification methods (SVM, decision trees, probabilistic models, nearest neighbors, etc...) Classification accuracy is improved by the used of superresolved images but, more importantly, so are the generalization ability of the classifiers. This is shown in the following two images, demonstrating the improvement in land recognition between the use of the original 500m MODIS data and the superresolved 250m data.

- A wavelet based resolution enhancement technique has been critically analyzed to see the effect of it on resolution enhancement modified discrete wavelet transform and interpolation based technique is proposed for enhancing the resolution of satellite images having low resolution in such a way that a highly resolved satellite image can be obtained without losing any image information. The advent of DWT has given a major impetus to many techniques based on achieving super resolution starting with a single low resolution image. In the proposed method, DWT is employed on the input satellite image to decompose it into sub-bands then the high frequency subbands and the input low resolution satellite image have been interpolated to obtain four interpolated images which are later combined after minor alterations to the interpolated input image using IDWT. The quantitative peak signal-to-noise ratio (PSNR) and classification results show that the resolution has been enhanced to a good scale without losing any information content present in the satellite image. The quality assessment parameters also illustrate the supremacy of the proposed technique over the conventional techniques. Results are shown in fig 8.

Figure 9. Classification MODIS image by developed approach.

- A technique based on feature extarction has been attempted to apply in the low resolution satellite data by which a land cover monitoring system can be developed. Moderate resolution imaging spectroradiometer (MODIS) data is a good resource for land cover monitoring as it is freely available data, having high temporal frequency and spatial resolutions 250 m, 500m and 1000m. MODIS being optical satellite data suffers from various atmospheric and cloud disturbances due to which, feature extraction and land cover interpretation using MODIS data is a significant and challenging task. In the past various features like spectral indices (EVI, SAVI, GEMI, PAVI etc), fourier based features,
wavelet based features were extracted for land cover classification from MODIS data but the role of texture descriptors and color features in land cover analysis has not been focused, which has the potential to provide a new stage of land cover discrimination. Therefore, the objective of this work was to explore the applicability of MODIS composite data for land cover monitoring by texture and color features extraction. Various texture features and descriptors like GLCM (Gray Level Co-occurrence Matrix) measures, LBP (Local Binary Pattern), EHD (Edge Histogram Descriptor), gabor wavelets and color features like Red-Green-Blue (RGB) color space, Hue-Saturation-Value (HSV) color space, Hue-Min-Max-Difference (HMMD) color space, MPEG-7 Dominant Color Descriptor (CD), MPEG-7 Color Structure Descriptor (CSD) and MPEG-7 Scalable Color Descriptor (SCD) were extracted. These color features were extracted over the artificial-color image obtained by mapping band2 (infrared band), NDVI (Normalized Vegetation Index) and band1 (red band) to the red, green and blue (RGB) color channels, respectively. It is observed that the extracted features are giving quite good results for land cover identification and classification. This infers that in near future these features could play a major role in the development of the land monitoring system using MODIS data. A classified result of Roorkee region of India is shown in fig 9 which has the overall classification accuracy approx. 82%.

- An another approach based on KLT (Kanade-Lucas-Tomasi) tracker has also been explored to apply on the Phased Array L-Band Synthetic Aperture Radar (PALSAR) satellite image for adaptive monitoring the land cover changes. It is observed that KLT tracking algorithm has good potential to be used as monitoring of vegetation in less time without applying time consuming image registration technique.

Related publications: link to list of publications on OPTIC web site.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific events organisation

10.1.1.1. General chair, scientific chair

N. Brodu is the principal convenor of the special session « Machine Learning adaptations for Earth monitoring » at European Geosciences Union General Assembly, the most important european congress in Earth Sciences.

10.1.1.2. Member of the conference program committees

H. Yahia was a member of the conference program committee Recent Advances in Electronics & Computer Engineering (RAECE), January 2015.

10.1.2. Journal

10.1.2.1. Member of the editorial boards

H. Yahia is a member of the editorial board of the open access journal Frontiers in Fractal Physiology.

10.1.3. Invited talks

- B. Xu has given an invited talk at the University of Basel, Research Group : Computational Physiology and Biostatistics. Title: From the Complexity Analysis of Biosignals to Clinical Applications.
- K. Daoudi has given 2 invited talks on nonlinear speech processing at the Czech Technical University of Prague and the Brno University (Czech Republic).
- H. Badri has given an oral presentation at the ORASIS conference [27].
O. Pont has given a presentation at the **SCAM seminar**: *Microcanonical cascade processes: how singularity analysis characterizes cardiac arrhythmia*, on November 19th, 2015.

### 10.1.4. Scientific expertise

- H. Yahia is a member of CNU (Conseil National des Universités), section 61.
- H. Yahia has participated in the evaluation of an ANR project.

### 10.2. Teaching - Supervision - Juries

#### 10.2.1. Teaching

- **Licence**: Hicham Badri, Unix Shell scripting & Python, 32 hours, L2 level, Bordeaux 1 University, France
- **Master**: Khalid Daoudi, Financial mathematics, 20 hours lecture courses, M2 level, Lorraine University, France
- **3rd year engineer school**: Nicoals Brodu, 18 hours, M1 level (supervision of 4 engineer students), Institut d’Optique, Bordeaux, France

#### 10.2.2. Supervision

- **PhD**: Hicham Badri, Sparse and Scale-Invariant Methods in Image Processing, Bordeaux 1 University, December 1st 2015, supervisors: H. Yahia and D. Aboutajdine, [14].
- **PhD**: Ayoub Tamim, Segmentation et classification des images satellitaires : application à la détection des zones d’upwelling côtier marocain et mise en place d’un applicatif de suivi spatio-temporel, Rabat University, September 22 2015, supervisors: K.Daoudi, H. Yahia, D. Aboutajdine, HAL link.
- **PhD in progress**: Camila Artana, Ocean dynamics at super-resolution Western Atlantic, defense scheduled in 2018, supervision: C. Provost and H. Yahia.
- **PhD in progress**: Akanksha Garg, super-resolution for Earth Observation, defense scheduled in 2017, co-supervision: N. Brodu and D. Singh (in the framework of the OPTIC associated team).
- **PhD in progress**: Ghopal Singh, Novel methods in classification and machine learning for Earth Observation, co-supervision: N. Brodu and D. Singh (in the framework of the OPTIC associated team).

#### 10.2.3. Juries

- H. Yahia is a member of the jury in the PhD defense of Mr. M. Osadebey, title *Noise Estimation, Noise Reduction and Intensity Inhomogeneity Correction in MRI Images of the Brain*, Concordia University, Electrical and Computer Engineering, June 15th, 2015.
- H. Yahia was asked by Bourgogne University to review HDR candidacy.

### 10.3. Popularization

Geostat has participated to the Inria initiative for the preparation of the COP21 international conference: *Our common future under climate change*.

### 11. Bibliography

**Major publications by the team in recent years**

[1] H. Badri, H. Yahia, K. Daoudi. *Fast and Accurate Texture Recognition with Multilayer Convolution and Multifractal Analysis*, in "European Conference on Computer Vision", Zürich, Switzerland, ECCV 2014, September 2014, [https://hal.inria.fr/hal-01064793](https://hal.inria.fr/hal-01064793)
[2] I. HERNÁNDEZ-CARRASCO, J. SUDRE, V. GARÇON, H. YAHIA, C. GARBE, A. PAULMIER, B. DEWITTE, S. ILLIG, I. DADOU, M. GONZÁLEZ-DÁVILA, J. SANTANA CASIANO. Reconstruction of super-resolution ocean pCO2 and air-sea fluxes of CO2 from satellite imagery in the Southeastern Atlantic, in "Biogeo- sciences", September 2015, 20 p. This work is a contribution to ESA Support To Science Element Grant N 400014715/11/I-NB OceanFlux- Upwelling Theme. The Surface Ocean CO2 Atlas (SOCAT) is an international effort, supported by the International Ocean Carbon Coordination Project (IOCCP), the Surface Ocean Lower Atmosphere Study (SOLAS), and the Integrated Marine Biogeochemistry and Ecosystem Research program (IMBER), to deliver a uniformly quality-controlled surface ocean CO2 database. The many researchers and funding agencies responsible for the collection of data and quality control are thanked for their contributions to SOCAT, https://hal.inria.fr/hal-01193339


**Publications of the year**

**Doctoral Dissertations and Habilitation Theses**


**Articles in International Peer-Reviewed Journals**


[19] I. Hernández-Carrasco, J. Sudre, V. Garçon, H. Yahia, C. Garbe, A. Paulmier, B. Dewitte, S. Illig, I. Dadou, M. González-Dávila, J. Santana Casiano. *Reconstruction of super-resolution ocean pCO 2 and air-sea fluxes of CO 2 from satellite imagery in the Southeastern Atlantic*, in "Biogeosciences", September 2015, 20 p. . This work is a contribution to ESA Support To Science Element Grant N 400014715/11/I-NB OceanFlux- Upwelling Theme. The Surface Ocean CO2 Atlas (SOCAT) is an international effort, supported by the International Ocean Carbon Coordination Project (IOCCP), the Surface Ocean Lower Atmosphere Study (SOLAS), and the Integrated Marine Biogeochemistry- istry and Ecosystem Research program (IMBER), to deliver a uniformly quality-controlled surface ocean CO2 database. The many researchers and funding agencies responsible for the collection of data and quality control are thanked for their contributions to SO- CAT, https://hal.inria.fr/hal-01193339


Invited Conferences


International Conferences with Proceedings


[27] H. BADRI, H. YAHIA. Régularisation parcimonieuse pour le problème d’intégration en traitement d’images, in "Journées francophones des jeunes chercheurs en vision par ordinateur", Amiens, France, June 2015, 6 p. , https://hal.archives-ouvertes.fr/hal-01161846


[29] O. YELEKCI, G. CHARRIA, X. CAPET, G. REVERDIN, J. SUDRE, H. YAHIA. Spatial and Seasonal Distributions of Frontal Activity over the French continental shelf in the Bay of Biscay observed from satellite Sea Surface Temperature, in "European Geosciences Union General Assembly 2015", Vienna, Austria, E. G. UNION (editor), Open Session on Coastal and Shelf Seas, EGU, April 2015, vol. OS2.1, https://hal.inria.fr/hal-01120612

Conferences without Proceedings

[31] A. JAYA KUMAR, K. DAoudi. Discrimination between pathological voice categories using matching pursuit, in "IEEE International Work Conference on Bioinspired Intelligence (IWOBI)", San Sebastian, Spain, June 2015 [DOI : 10.1109/IWOBI.2015.7160169], https://hal.inria.fr/hal-01250441

Scientific Books (or Scientific Book chapters)


Other Publications

[33] G. Attuel, H. Yahia. The role of abnormal inhibitory transmission at the gap junctions of cardiac cells in fibrillation, January 2016, Physical Principles of Biological and Active Systems - IOP Topical Meeting, Poster, https://hal.inria.fr/hal-01253790

[34] O. Pont, A. Turiel, H. Yahia. A method for computing singularity exponents by means of reconstructability measures, 2015, SIAM Multiscale Modeling and Simulation, submitted, https://hal.inria.fr/hal-00688717

[35] H. Yahia, J. Sudre, I. Hernández-Carrasco, V. Garçon, D. Singh, N. Brodu, C. Garbe. General methodology for the derivation of high resolution oceanic data through information fusion at different scales, October 2015, working paper or preprint, https://hal.inria.fr/hal-01217596

References in notes


