



PANAMA

Parsimony & New Algorithms for Signal & Audio Modeling

Inria – Rennes Bretagne Atlantique

Theme: Language, Audio, and Speech

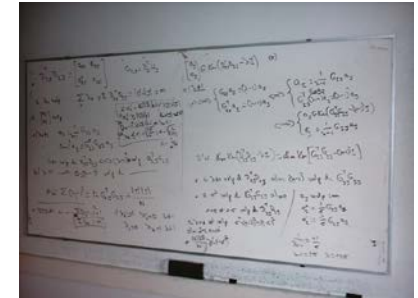
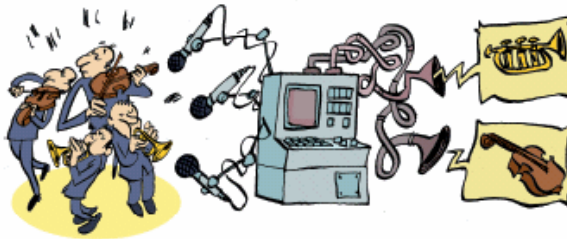
IRISA – CNRS UMR 6074

Dept D5 - Digital Signals and Images, Robotics

Présentation : Frédéric BIMBOT

Equipe PANAMA - Inria-Rennes/IRISA

- ***Parcimonie et Nouveaux Algorithmes pour le Signal et la Modélisation Audio***



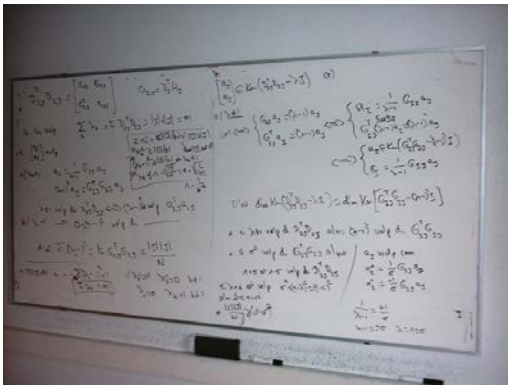
- **Composition**

- ✓ 4 permanents (2 INRIA + 2 CNRS)
- ✓ Total ~ 20 personnes (doctorants, postdocs, ingénieurs)
- ✓ 15 ans d'expérience acquise dans le domaine
- ✓ Responsable scientifique : Rémi Gribonval (Inria)

www.irisa.fr/panama



● *Traitement du signal pour l'audio*



- ✓ Analyse et traitement des signaux audio, parlés, musicaux
- ✓ Localisation et séparation de sources sonores
- ✓ Représentations parcimonieuses et modèles structurels
- ✓ Réduction de dimension, acquisition et apprentissage compressif

Visées applicatives

● Expertise prédominante en audio ...

- ✓ **Capture, production and ré-éditorialisation** de contenus audio
- ✓ **Navigation et recherche d'information** dans les contenus audio(visuels)
- ✓ **Description** multi-niveaux et **exploitation** versatile de contenus musicaux
- ✓ **Education, loisir, divertissement** (jeux et jouets,...)
- ✓ **Restauration** audio, spatialisation, surveillance et monitoring sonore

● ... mais pas uniquement

À travers des collaborations

- ✓ Problèmes inverses en acoustique, apprentissage compressif
- ✓ Imagerie acoustique (classique et sous-marine)
- ✓ Analyse de signaux biomédicaux (ECG, EEG, multimodal ...)
- ✓ Audition robotique

Périmètre scientifique

Principaux fondements

Mathématiques et traitement du signal

Modélisation probabiliste et computationnelle

Conception d'algorithmes rapides et efficaces (critères et optimisation)

Evaluation expérimentale

Challenges

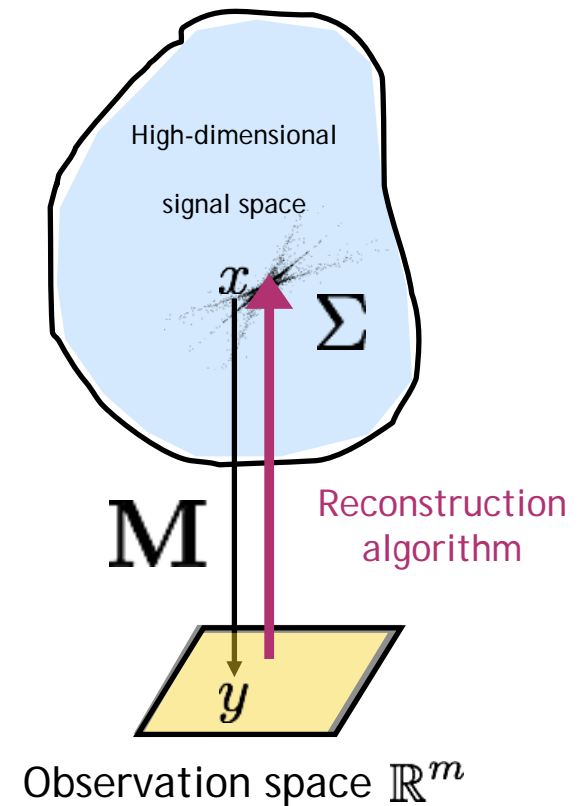
✓ **hétérogénéité**
parole
musique
autre

✓ **volume & dimension**
collections
flux
multi-captation

✓ **versatilité**
captation
délinéarisation
structuration
navigation

Problèmes inverses & réduction de dimension

- ✓ Projection en basse dimension
 - ◆ Représentations compactes
- ✓ Algorithmes de reconstruction
 - ◆ Séparation de sources
 - ◆ Restauration audio
- ✓ Inférence de modèle
 - ◆ Construction de dictionnaire
 - ◆ Découverte de motifs
 - ◆ Apprentissage compressif
- ✓ Calibration
 - ◆ Localisation de sources
 - ◆ Localisation de capteurs



Flexible Audio Source Separation

Fueled by research

FASST v2.0

- Flexibility (XML params)
- C++ core functions
- Matlab/Python wrappers
- A||GO web-service platform
- QPL licence
- **600 downloads**
with MULTISPEECH team

Feeding industrial collaborations

- ◆ *IEEE 2012 SPS Young Author **Best Paper Award***
Full-rank models of spatial covariance
- ◆ **IEEE SPM 2014 review paper**
Guided source separation

- ◆ Canon Research France
(2 patents)
- ◆ MAIA Studios
- ◆ VoiceHome (onMobile)

Fouille et structuration de contenus audio massifs

Découverte, détection et exploitation des répétitions, similarités et motifs

dans les contenus audio (flux d'information, musique, environnement sonore, etc...)

- **Réutilisation de contenus** : résumé, démixage, ...
- **Gestion de données** : mise en relation de concepts intra et inter collection
- **Interprétation de scènes sonores** : détection d'alarme, homologues musicales...

→ Motif discovery in audio data

✓ Ph.D. Souviraà-Labastie (2012-2015)

with LINKMEDIA team

Modélisation de la structure musicale

- **Structure sémiotique : description, annotation & détection automatique**

- ✓ Ph.D. Sargent (2010-2013)

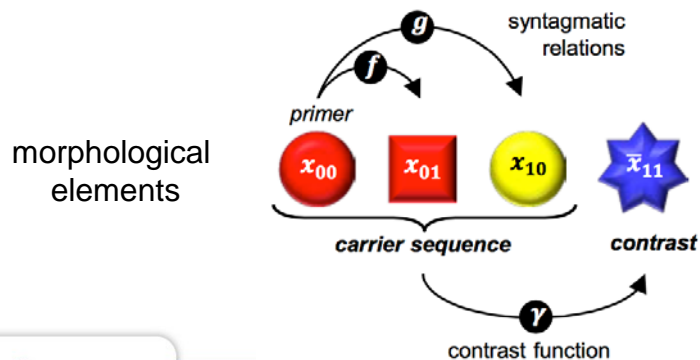
I/A A₁ A₂ B C J/2 A'₁ |A'₂ B' C X/C Y/2 C*

- ◆ Quaero/Metiss database
→ 380 annotated pop-songs
- ◆ 3 participations to MIREX evaluation campaign

- **Le modèle « Système-Contraste »**

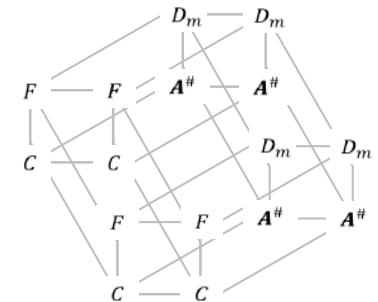
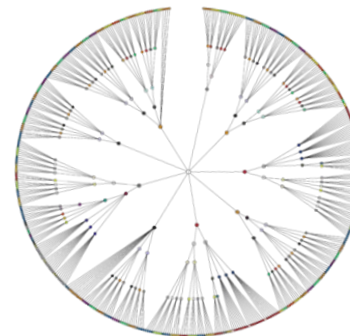
- ✓ Ph.D. Sargent (2010-2013), Master Deruty (2013)

→ *Méta-modèle intuitif de narration au sein d'un segment musical*



- **Tenseurs et graphes pour modéliser la structure musicale**

- ✓ Ph.D. Guichaoua (2012-...)
 - ✓ Ph.D. Louboutin (2015-...)
 - ✓ Collab. with P. Vanderghyest, EPFL



Hot topic !

Traitement de signal sur graphes

- **Rapidly emerging:**
Graph Signal Processing

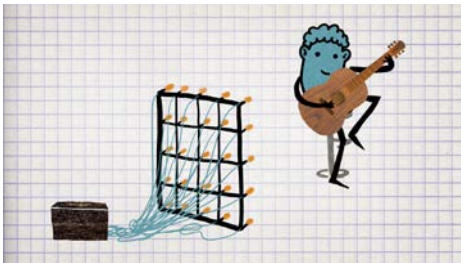
- ✓ ***Goal: to extend to signals on graphs many concepts from signal processing***
 - ◆ *Graph wavelets*
 - ◆ *Fast Transforms on Graphs*
 - ◆ *Compressive Spectral Clustering*
- ✓ **Chaire Inria P. Vandergheynst**
2014-...
- ✓ **EPFL/Inria lab**
2015-...

Vulgarisation

-Pour La Science magazine



-Interstice web site video



5 Journées « Science & Musique »

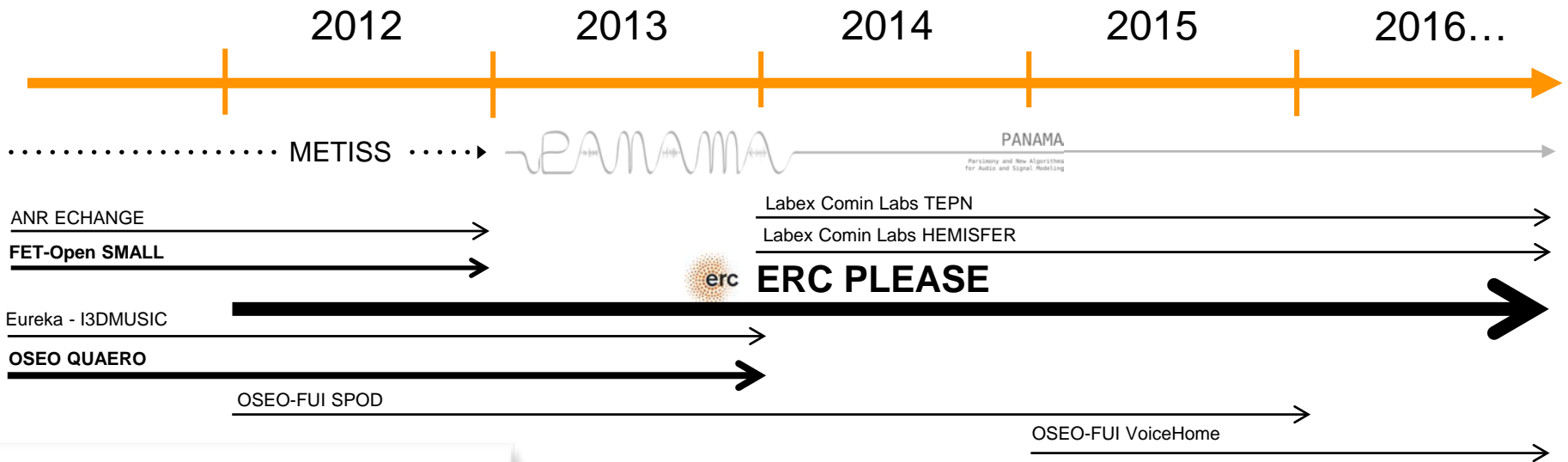
2011-2015: plus de 250 participants chaque année



Merci de votre attention

Contact : frederic.bimbot@irisa.fr

Funding and collaborations



Sparse Models & Representations

**EPFL (CH),
Univ. Edimbourg (UK)
Queen Mary University (UK)
Technion (ISR)**

SIERRA team (Inria Paris)

TU Munich (DE)
Univ. Postdam (DE)

Institut Langevin (UPMC)
LATP/LIF (Aix-Marseille Univ.)

Technicolor (Rennes)

Acoustic Scene Analysis

**Institut Langevin (UPMC)
MULTISPEECH team (Inria Nancy)
LAGADIC team (Inria Rennes)**

LATP/LIF (Aix-Marseille Univ.)
Télécom Paris Tech, IRCAM

**Canon Research France (Rennes)
Studio MAIA (Paris)
OnMobile SA (Rennes)
DeltaDore (Rennes)
Orange Labs (Rennes)
Audionamix (Paris)**

Audio Mining

**LINKMEDIA team (Inria Rennes)
MULTISPEECH team (Inria Nancy)**

Studio MAIA (Paris)

Brain imaging & Gesture

**VISAGES team (Inria Rennes)
HYBRID team (Inria Rennes)
LTSI (Univ. Rennes I)**

(Data Jungle) x (Data Deluge)

- **Data Jungle**

- ✓ Complex content — beyond handmade / expert modeling

- **Data Deluge**

- ✓ High dimensional items (ex: massively multichannel)

- ✓ Large collections / stream (ex: audiovisual)

- Opportunity: **latent information**

Learning from huge collections & streams

Multiscale **structuring**

Drastic **dimension reduction**

Research Axes

foundations of signal processing and machine learning

- **Drastic dimension-reduction**

- ◆ tractable low-dimensional models
- ◆ scalable & distributed algorithms
- ◆ provably good sketched learning
- ◆ foundations of graph signal processing

Graph
Signal Processing

from signal level to semantic level

- **Multi-scale structuring
of audio, speech & music**

- ◆ contents and streams
- ◆ inner structure of objects and relations between objects
- ◆ graphs and tensors

Acoustic
Space
Learning

versatile approaches for efficient sensing

- **Large-scale & robust
acoustic learning**

- ◆ large ad-hoc sensor arrays
- ◆ mobile & robot audition
- ◆ multilayer latent acoustic models
- ◆ distributed acoustic sensing

Sketched
Learning