Multi-componets Data, Signal and Image Processing for Biological and Medical Applications

Ali Mohammad-Djafari

Laboratoire des Signaux et Systèmes
UMR 8506 CNRS - CS - Univ Paris Sud
CentraleSupélec, Gif-sur-Yvette.

djafari@lss.supelec.fr
http://djafari.free.fr

January 6, 2017

Summary 2.1: Time series

- Time serie and Fourier representation
  - Continuous / Discrete
  - Correlation, Inter-correlation, Inter-dependance
  - Stationarity / non-stationarity
  - Convolution and Deconvolution
- Filtering and Denoising
- Modelling and Prediction
- Parametric and Non parametric models
- Parametric models:
  - Least Squares
  - Maximum Likehood
  - Bayesian estimation

Summary 2.2: Images

- Continuous / Discrete
- Gray and Color images
  - 2D FT and FFT
  - 2D Correlation and inter-correlation
  - Stationarity / non-stationarity
  - 2D Convolution
- Filtering and Denoising
- Modelling and Prediction
- Simple Markovian models
- Contours and Regions
- Hierarchical Markov models

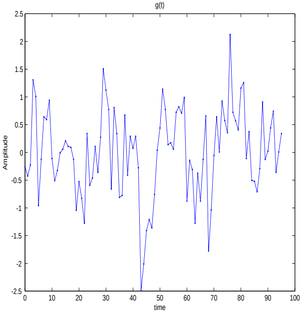
1D signals: Time series

- 1D Signal: Time series: $x_i = f(t_i)$
- In general no exchangeable.
- Time representation $f(t)$, Fourier Transform and Fourier representation $F(\nu)$, Auto Correlation function $R(\tau)$, Power Spectral Density $S(\nu)$
- Stationary and non stationary
- STFT, Time-Frequency, Time-Scale, Wavelets, ...
- Smoothing, Noise removing, Filtering
- Periodic signals, estimation of the period, Fourier series
- Modeling:
  - Sum of sinusoids model and parameter estimation
  - Moving average (MA) model
  - Autoregressive (AR) model

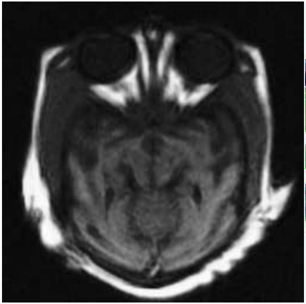
Representation of signals and images

- Signal: $f(t), f(x), f(\nu)$
  - $f(t)$ Variation of temperature in a given position as a function of time t
  - $f(x)$ Variation of temperature as a function of the position x on a line
  - $f(\nu)$ Variation of temperature as a function of the frequency ν
- Image: $f(x,y), f(x,t), f(\nu,t), f(\nu_1,\nu_2)$
  - $f(x,y)$ Distribution of temperature as a function of the position (x,y)
  - $f(x,t)$ Variation of temperature as a function of x and t
  - ...
- 3D, 3D+t, 3D+ν, ... signals: $f(x,y,z), f(x,y,t), f(x,y,z,t)$
  - $f(x,y,z)$ Distribution of temperature as a function of the position (x,y,z)
  - $f(x,y,z,t)$ Variation of temperature as a function of (x,y,z) and t
  - ...

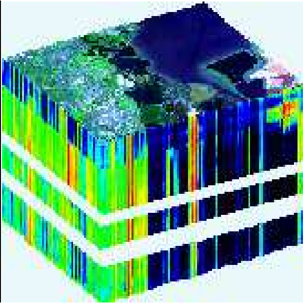
Representation of signals



1D signal



2D signal=image



3D signal

Linear Transformations

$$g(\boldsymbol{s}) = \int_D f(\boldsymbol{r})\, h(\boldsymbol{r}, \boldsymbol{s})\, \mathrm{d}\boldsymbol{r}$$

$$f(\boldsymbol{r}) \longrightarrow \boxed{h(\boldsymbol{r}, \boldsymbol{s})} \longrightarrow g(\boldsymbol{s})$$

- 1–D :

$$g(t) = \int_D f(t')\, h(t, t')\, \mathrm{d}t'$$

$$g(x) = \int_D f(x')\, h(x, x')\, \mathrm{d}x'$$

- 2–D :

$$g(x, y) = \iint_D f(x', y')\, h(x, y; x', y')\, \mathrm{d}x'\, \mathrm{d}y'$$

$$g(r, \phi) = \iint_D f(x, y)\, h(x, y; r, \phi)\, \mathrm{d}x\, \mathrm{d}y$$

Linear and Invariant systems: convolution

$$h(\boldsymbol{r}, \boldsymbol{r}') = h(\boldsymbol{r} - \boldsymbol{r}')$$

$$f(\boldsymbol{r}) \longrightarrow \boxed{h(\boldsymbol{r})} \longrightarrow g(\boldsymbol{r}) = h(\boldsymbol{r}) * f(\boldsymbol{r})$$

- 1–D :

$$g(t) = \int_D f(t')\, h(t - t')\, \mathrm{d}t'$$

$$g(x) = \int_D f(x')\, h(x - x')\, \mathrm{d}x'$$

- 2–D :

$$g(x, y) = \iint_D f(x, y)\, h(x - x', y - y')\, \mathrm{d}x'\, \mathrm{d}y'$$

- $h(t)$ impulse response
- $h(x, y)$ Point Spread Function

Linear Transformations: Separable systems

$$g(\boldsymbol{s}) = \int_D f(\boldsymbol{r})\, h(\boldsymbol{r}, \boldsymbol{s})\, \mathrm{d}\boldsymbol{r}$$

$$h(\boldsymbol{r}, \boldsymbol{s}) = \prod_j h_j(r_j, s_j)$$

Examples:

- 2D Fourier Transform

$$g(\omega_x, \omega_y) = \iint f(x, y)\, \exp\left[-j(\omega_x x + \omega_y y)\right]\, \mathrm{d}x\, \mathrm{d}y$$

$$h(x, y, \omega_x, \omega_y) = h_1(\omega_x x)\, h_2(\omega_y y)$$

$$\exp\left[-j(\omega_x x + \omega_y y)\right] = \exp\left[-j(\omega_x x)\right]\, \exp\left[-j(\omega_y y)\right]$$

- nD Fourier Transform

$$g(\omega) = \int f(\boldsymbol{x})\, \exp\left[-j\omega' \boldsymbol{x})\right]\, \mathrm{d}\boldsymbol{x}$$

Fourier Transform

[Joseph Fourier, French Mathematicien (1768-1830)]

- 1D Fourier: $\mathcal{F}_1$

$$\left\{\begin{array}{ll} g(\omega) & = \int f(t) \exp\left[-j\omega t\right] \,\mathrm{d}t \\ f(t) & = \frac{1}{2\pi} \int g(\omega) \exp\left[+j\omega t\right] \,\mathrm{d}\omega \end{array}\right.$$

- 2D Fourier: $\mathcal{F}_2$

$$\left\{\begin{array}{ll} g(\omega_x, \omega_y) & = \iint f(x, y) \exp\left[-j(\omega_x x + \omega_y y\right] \,\mathrm{d}x \,\mathrm{d}y \\ f(x, y) & = (\frac{1}{2\pi})^2 \iint g(\omega_x, \omega_y) \exp\left[+j(\omega_x x + \omega_y y\right] \,\mathrm{d}\omega_x \,\mathrm{d}\omega_y \end{array}\right.$$

- nD Fourier: $\mathcal{F}_n$

$$\left\{\begin{array}{ll} g(\boldsymbol{\omega}) & = \int f(\boldsymbol{x}) \exp\left[-j\boldsymbol{\omega}'\boldsymbol{x}\right] \,\mathrm{d}\boldsymbol{x} \\ f(\boldsymbol{x}) & = (\frac{1}{2\pi})^n \int g(\boldsymbol{\omega}) \exp\left[+j\boldsymbol{\omega}'\boldsymbol{x}\right] \,\mathrm{d}\boldsymbol{\omega} \end{array}\right.$$

1D Fourier Transform $\mathcal{F}_1$

$$\left\{\begin{array}{ll} g(\omega) & = \int f(t) \exp\left[-j\omega t\right] \mathrm{d}t \\ f(t) & = \frac{1}{2\pi} \int g(\omega) \exp\left[+j\omega t\right] \mathrm{d}\omega \end{array}\right.$$

- $|g(\omega)|^2$ is called the spectrum of the signal $f(t)$
- For real valued signals $f(t)$, $|g(\omega)|$ is symetric

Examples:

$f(t)$	$g(\omega)$
$\exp\left[-j\omega_0 t\right]$	$\delta(\omega-\omega_0)$
$\sin(\omega_0 t)$	?
$\cos(\omega_0 t)$	?
$\exp\left[-t^2\right]$	?
$\exp\left[-\frac{1}{2}(t-m)^2/\sigma^2\right]$	?
$\exp\left[-t/\tau\right], t>0$	?
$1 \text{ if } \|t\| < T/2$	?

2D Fourier Transform: $\mathcal{F}_2$

$$
\begin{cases}
g(\omega_x,\omega_y) & = \iint f(x,y) \exp\left[-j(\omega_x x + \omega_y y\right] \, \mathrm{d}x \, \mathrm{d}y \\
f(x,y) & = (\frac{1}{2\pi})^2 \iint g(\omega_x,\omega_y) \exp\left[+j(\omega_x x + \omega_y y\right] \, \mathrm{d}\omega_x \, \mathrm{d}\omega_y
\end{cases}
$$

- $|g(\omega_x,\omega_y)|^2$ is called the spectrum of the image $f(x,y)$
- For real valued image $f(x,y)$, $|g(\omega_x,\omega_y)|$ is symetric with respect of the two axis ω_x and ω_y.

Examples:

$f(x,y)$	$g(\omega_x,\omega_y)$
$\exp\left[-j(\omega_{x0}x + \omega_{y0}y)\right]$	$\delta(\omega_x - \omega_{x0})\delta(\omega_y - \omega_{y0})$
$\exp\left[-(x^2+y^2)\right]$	?
$\exp\left[-\frac{1}{2}[(x-m_x)^2/\sigma_x^2 + (y-m_y)^2/\sigma_y^2]\right]$	?
$\exp\left[-(\lvert x\rvert + \lvert y\rvert)\right]$	?
1 if $\lvert x\rvert < T_x/2$ & $\lvert y\rvert < T_y/2$	?
1 if $(x^2+y^2) < a$	?

nD Fourier Transform: $\mathcal{F}_n$

$$\left\{\begin{array}{ll} g(\omega) & = \int f(\boldsymbol{x}) \exp\left[-j\omega'\boldsymbol{x}\right] \, \mathrm{d}\boldsymbol{x} \\ f(\boldsymbol{x}) & = (\frac{1}{2\pi})^n \int g(\omega) \exp\left[+j\omega'\boldsymbol{x}\right] \, \mathrm{d}\omega \end{array}\right.$$

- $|g(\omega)|^2$ is called the spectrum of $f(\boldsymbol{x})$
- For real valued image $f(\boldsymbol{x})$, $|g(\omega)|$ is symetric with respect of all the axis ω_j.

Examples:

$f(\boldsymbol{x})$	$g(\omega)$
$\exp\left[-j(\omega_0'\boldsymbol{x})\right]$	$\delta(\omega-\omega_0)$
$\exp\left[-\boldsymbol{x}'\boldsymbol{x}\right] = \exp\left[-\|\boldsymbol{x}\|^2\right]$	$\exp\left[-\omega'\omega\right] = \exp\left[-\|\omega\|^2\right]$
$\exp\left[-\|\boldsymbol{D}\boldsymbol{x}\|^2\right]$	?
$\exp\left[-\frac{1}{2}[(\boldsymbol{x}-\boldsymbol{m})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{m})\right]$	?
1 if $\|x\|^2 < R$	?
1 if $|x_j| < R$	?

Hilbert Transform: $\mathcal{H}$

[David Hilbert, German mathematicien (1862-1943)]

- Definition: If $f \in L_2$ on $(-\infty, \infty)$,

$$g(x) = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{f(t)}{t-x} \, \mathrm{d}t$$

$$f(t) = \frac{-1}{\pi} \int_{-\infty}^{+\infty} \frac{g(x)}{x-t} \, \mathrm{d}x$$

  The integrals are interpreted in the Cauchy principal value(CPV) sense at $t = x$.

- Alternate expression useful in signal processing:

$$g(t) = \frac{1}{\pi} \lim_{\epsilon \mapsto 0} \int_{\epsilon}^{\infty} \frac{f(t+\tau) - f(t-\tau)}{\tau} \, \mathrm{d}\tau$$

Hilbert Transform: $\mathcal{H}$

- If $f \in L_2$
  - $\mathcal{H}\left(\mathcal{H}(f)\right) = f$
  - f and $\mathcal{H}(f)$ are orthogonal, i.e.,

$$\lim_{r\to\infty} \int_{-r}^{r} [f\mathcal{H}(f)](u)\,\mathrm{d}u = 0$$

  - The Hilbert transform of a constant is zero.
  - Hilbert and Fourier Transforms

$$\mathcal{H}(f) = f * \frac{-1}{\pi t} \quad \longrightarrow \mathcal{F}\{\mathcal{H}(f)\} = j\mathrm{sgn}(\omega) F(\omega)$$

Radon Transform (RT): $\mathcal{R}$

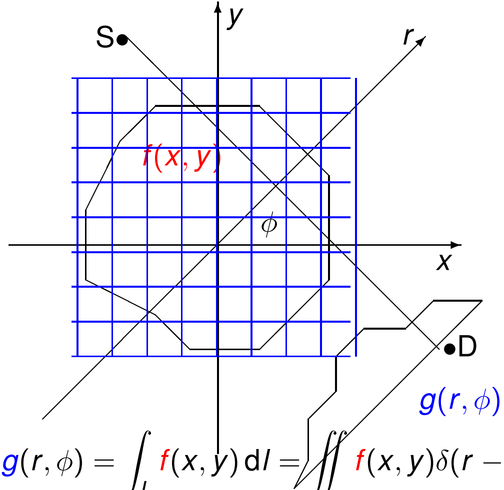
- Definition:
  This transform is defined for the functions in 2 or more dimensions. Here we give the relations only in the 2–D case.

$$
\begin{aligned}
g(r,\phi) &= \int_{L_{r,\phi}} f(x,y)\,\mathrm{d}l \\
&= \iint f(x,y)\,\delta(r - x\cos(\phi) - y\sin(\phi))\,\mathrm{d}x\,\mathrm{d}y
\end{aligned}
$$

- The Radon transform maps the spatial domain $(x,y) \in \mathrm{R}^2$ to the domain $(r,\phi) \in \mathrm{R} \times [0,\pi]$. Each point in the (r,ϕ) space corresponds to a line in the spatial domain (x,y).
- Note that (r,ϕ) are not the polar coordinates of (x,y). In fact if we note the polar coordinates corresponding to the (x,y) point (ρ,θ), then we have

$$
x = \rho\cos\theta, \quad y = \rho\sin\theta, \quad r = \rho\cos(\phi - \theta)
$$

X ray Tomography



$$I = I_0 \exp\left[-\int_L f(x)\,\mathrm{d}x\right]$$

$$-\ln\frac{I}{I_0} = \int_L f(x,y)\,\mathrm{d}l$$

$$g(r,\phi) = \int_L f(x,y)\,\mathrm{d}l$$

$$g(r,\phi) = \int_L f(x,y)\,\mathrm{d}l = \iint f(x,y)\delta(r - x\cos\phi - y\sin\phi)\,\mathrm{d}x\,\mathrm{d}y$$

$$f(x,y) \longrightarrow \boxed{\text{Radon Transform}} \longrightarrow g(r,\phi)$$

$$g(r,\phi) \longrightarrow \boxed{\text{Image Reconstruction}} \longrightarrow f(x,y)$$

Radon Transform: Some properties

[Johann K.A. Radon, Austrian mathematician (1887-1956)]

- Definition in cartezian coordinate system:

$$f(x,y) \xrightarrow{\mathcal{R}} g(r,\phi) = \iint f(x,y)\delta(r - x\cos(\phi) - y\sin(\phi))\,\mathrm{d}x\,\mathrm{d}y$$

- Definition in polar coordinate system:

$$f(\rho,\theta) \xrightarrow{\mathcal{R}} g(r,\phi) = \int_0^{\infty}\int_0^{2\pi} f(\rho,\theta)\delta(r - \rho\cos(\phi-\theta)\rho\,\mathrm{d}\rho\,\mathrm{d}\theta$$

- Inversion

$$f(x,y) = \frac{1}{2\pi}\int_0^{\pi}\int_0^{\infty} \frac{\partial g(r,\phi)/\partial r}{r - x\cos(\phi) - y\sin(\phi)}\,\mathrm{d}r\,\mathrm{d}\phi$$

Radon Transform: Inversion

Direct Inverse Radon Transform

$$g(r,\phi) \longrightarrow \boxed{\begin{array}{c}\text{Differentiate}\\ \frac{1}{2\pi}\mathcal{D}\end{array}} \longrightarrow \boxed{\begin{array}{c}\text{Hilbert Transform}\\ \mathcal{H}\end{array}} \xrightarrow{\widetilde{g}(r,\phi)} \boxed{\begin{array}{c}\text{Back–projection}\\ \mathcal{B}\end{array}} \xrightarrow{f(x,}$$

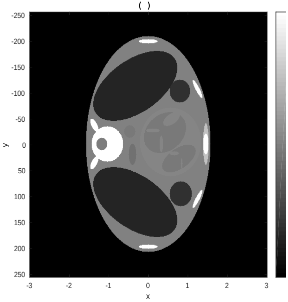
Convolution Back–projection method

$$g(r,\phi) \longrightarrow \boxed{\begin{array}{c}\text{1–D Filter}\\ |\Omega|\end{array}} \xrightarrow{\widetilde{g}(r,\phi)} \boxed{\begin{array}{c}\text{Back–projection}\\ \mathcal{B}\end{array}} \xrightarrow{f(x,y)}$$

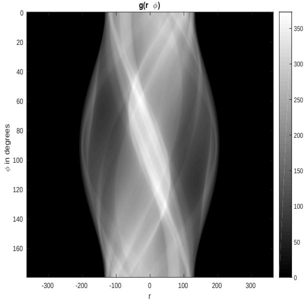
Filter Back–projection method

$$g(r,\phi) \longrightarrow \boxed{\begin{array}{c}\text{FT}\\ \mathcal{F}_1\end{array}} \longrightarrow \boxed{\begin{array}{c}\text{Filter}\\ \Omega\end{array}} \longrightarrow \boxed{\begin{array}{c}\text{IFT}\\ \mathcal{F}_1^{-1}\end{array}} \xrightarrow{\widetilde{g}(r,\phi)} \boxed{\begin{array}{c}\text{Back–projection}\\ \mathcal{B}\end{array}} \xrightarrow{f(x,y)}$$

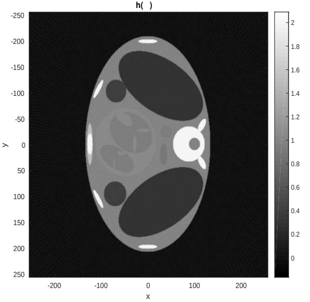
Radon Transform (RT) and Filtered Back-Projection (FBP) image reconstruction



a) original $f(x, y)$

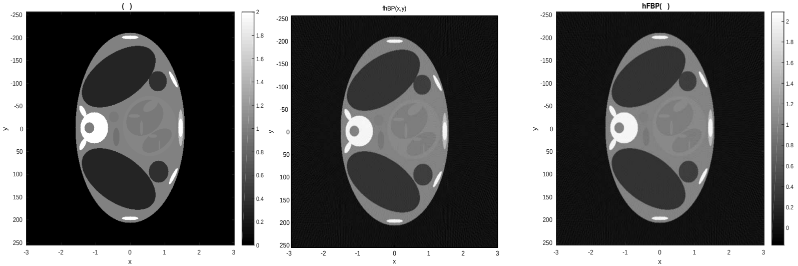


b) RT $g(r, \phi)$

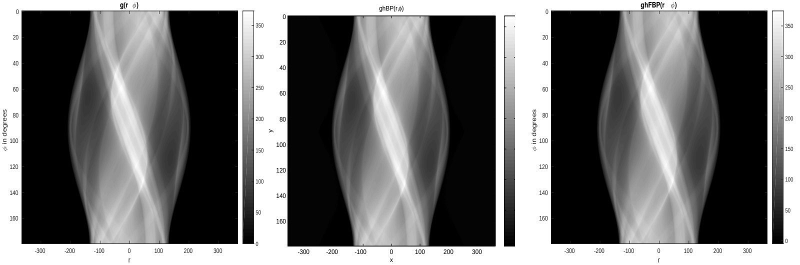


c) FBP reconstruction $\widehat{f}(x, y)$

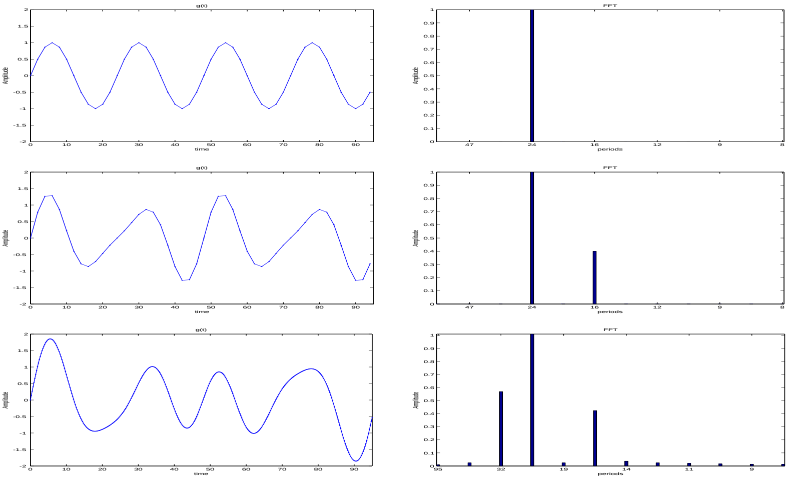
Radon Transform (RT), Back-Projection (BP) and Filtered Back-Projection (FBP) image reconstruction



a) $f(x,y)$ b) $\widehat{f}(x,y)$ by BP $\widehat{f}(x,y)$ by FBP



1D signals $f(t)$ and its FT $F(\nu)$



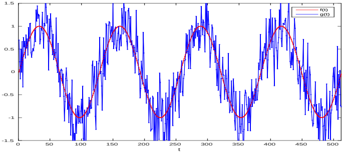
signal $f(t)$ Modulus of its FT $|F(\nu)|^2$

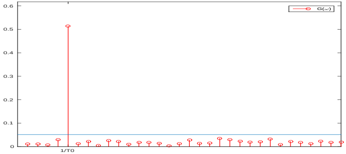
1D signals: Noise removing using FT

$$g(t) = f(t) + \epsilon(t)$$

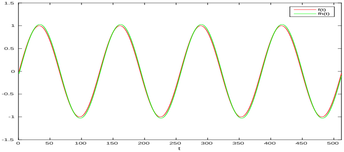
$$g(t) \longrightarrow \boxed{\text{FFT}} \longrightarrow \boxed{\text{Threshold}} \longrightarrow \boxed{\text{IFFT}} \longrightarrow \widehat{f}(t)$$

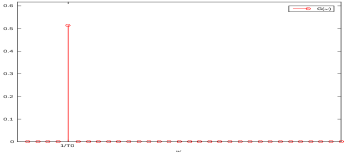
$g(t) = f(t) + \epsilon(t)$ →FFT→ $G(\omega) = F(\omega) + E(\omega)$





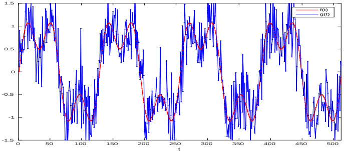
$\widehat{f}$ ←IFFT← Thresholded $G(\omega)$

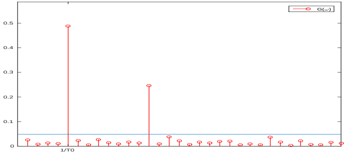




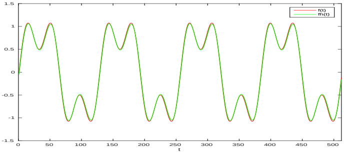
1D signals: Noise removing using FT

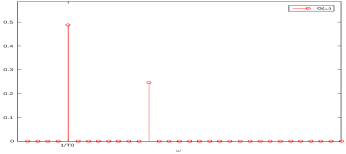
$g(t) = f(t) + \epsilon(t)$ $\quad \rightarrow$FFT$\rightarrow \quad$ $G(\omega) = F(\omega) + E(\omega)$





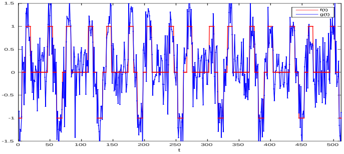
$\widehat{f}$ $\quad \leftarrow$IFFT$\leftarrow \quad$ Thresholded $G(\omega)$

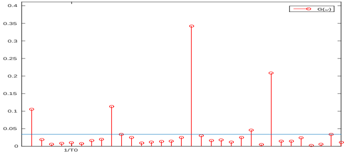




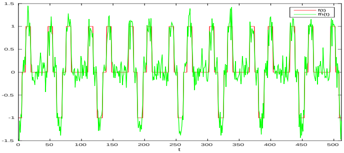
1D signals: Noise removing using FT

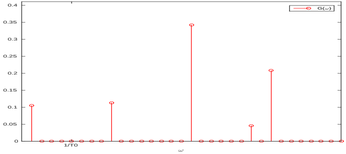
$$g(t) = f(t) + \epsilon(t) \quad \rightarrow \text{FFT} \rightarrow \quad G(\omega) = F(\omega) + E(\omega)$$



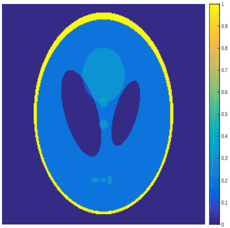


$$\widehat{f} \quad \leftarrow \text{IFFT} \leftarrow \quad \text{Thresholded } G(\omega)$$

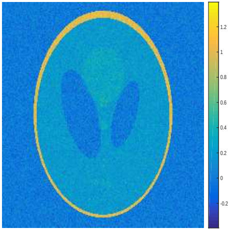




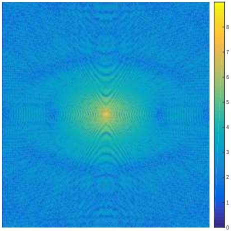
Images denoising using FFT



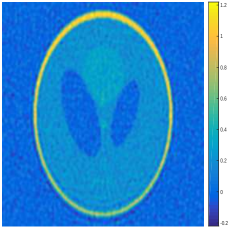
$f(x, y)$



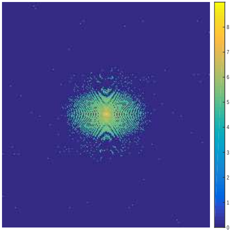
$g(x, y) = f(x, y) + \epsilon(x, y)$



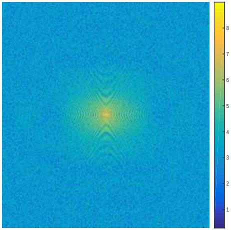
$F(u, v)$



$\widehat{f}(x, y)$



Thresholded $G(u, v)$

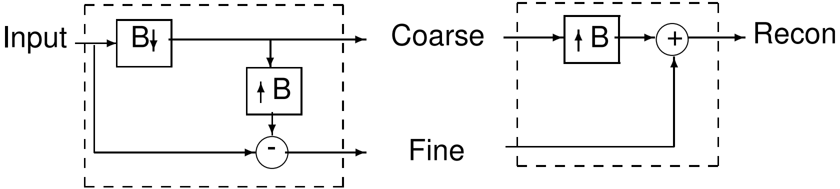


$G(u, v)$

Images: Pyramidal and Wavelet Transform

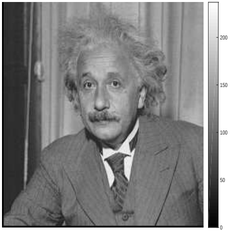
Pyramidal representation:

- Invertible Linear Transform
- image $\longrightarrow$ Coarse and Fine images
- B: Band Pass Filtering

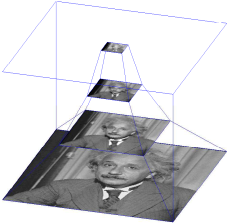


- Number of samples: $(1 + 1/2 + 1/4 + ...)$
- Overcomplete representation

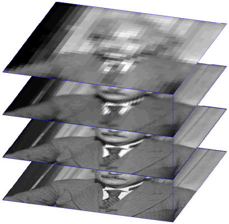
Images: Pyramidal representations



Original



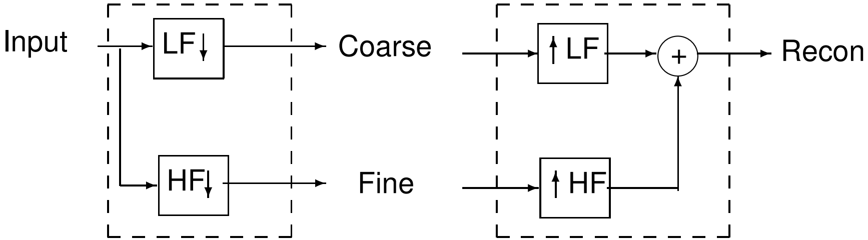
Pyramid



Pyramid

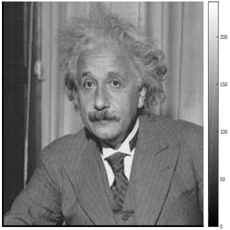
Images: QMF and Wavelet Transform

- Invertible Linear Transform
- image $\longrightarrow$ Coarse and Fine images
- Quadrature Mirror Filtering

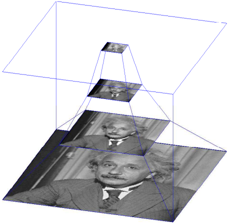


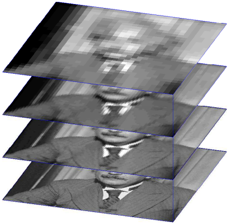
- Number of samples: 1
- Complete representation

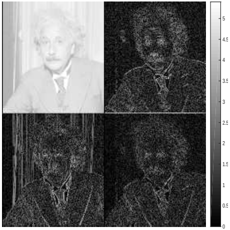
Images: Pyramidal representations

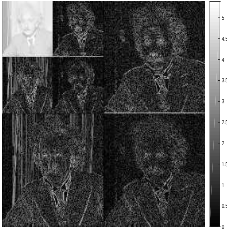


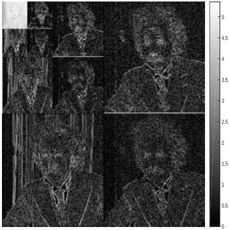
Original


Pyramid

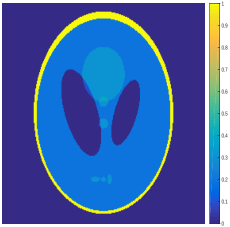

Pyramid


WT 1

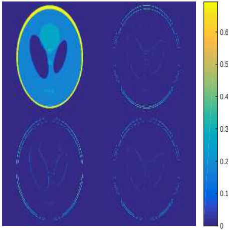

WT 2


WT 3

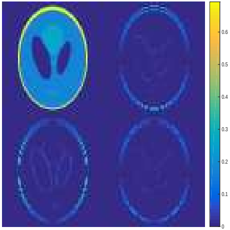
Images compression using WT



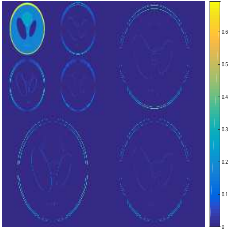
Original



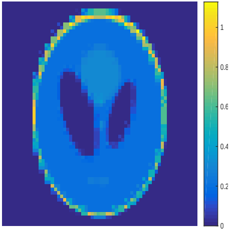
WT 1



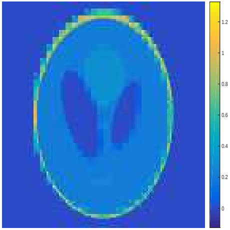
WT 2



WT 1-2

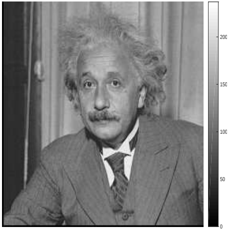


IWT 1-1

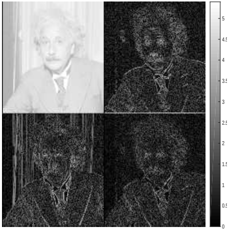


IWT 1-2

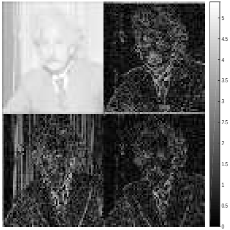
Images compression using WT



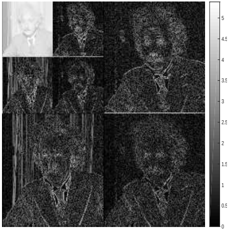
Original



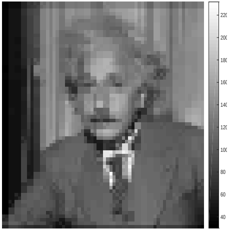
WT 1



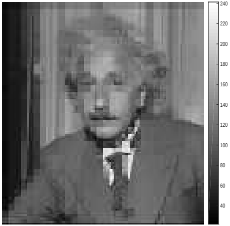
WT 2



WT 1-2

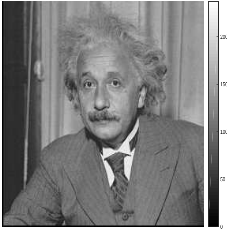


IWT 1-1

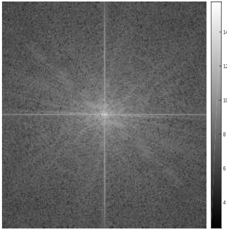


IWT 1-2

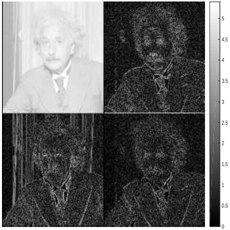
Images: Different representations



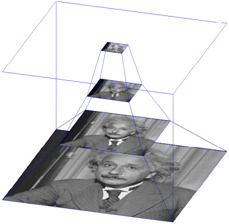
Original



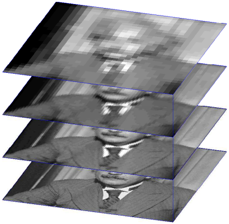
FFT



WT 2



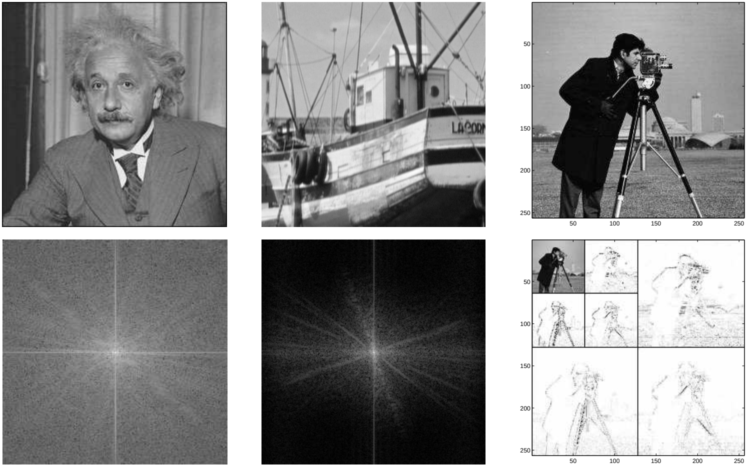
Pyramid



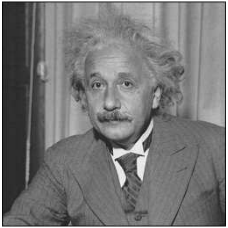
Pyramid

Laplacien

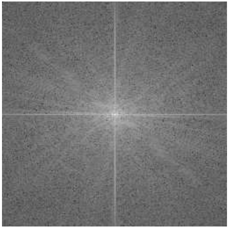
Images: Space, Fourier and Wavelets representations



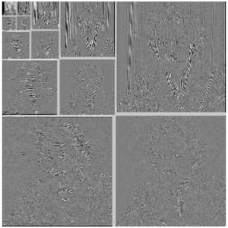
Sparse images (Fourier and Wavelets domain)



Image



Fourier



Wavelets

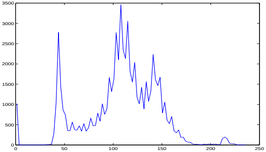
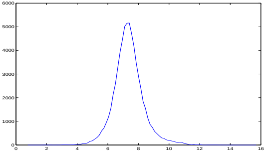
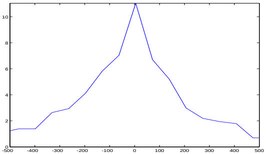


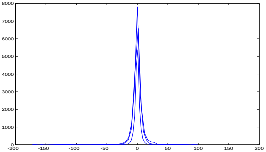
Image hist.



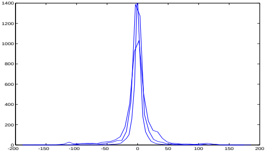
Fourier coeff. hist.



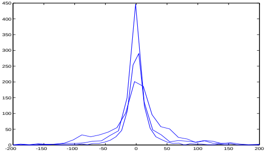
Wavelet coeff. hist.



bands 1-3

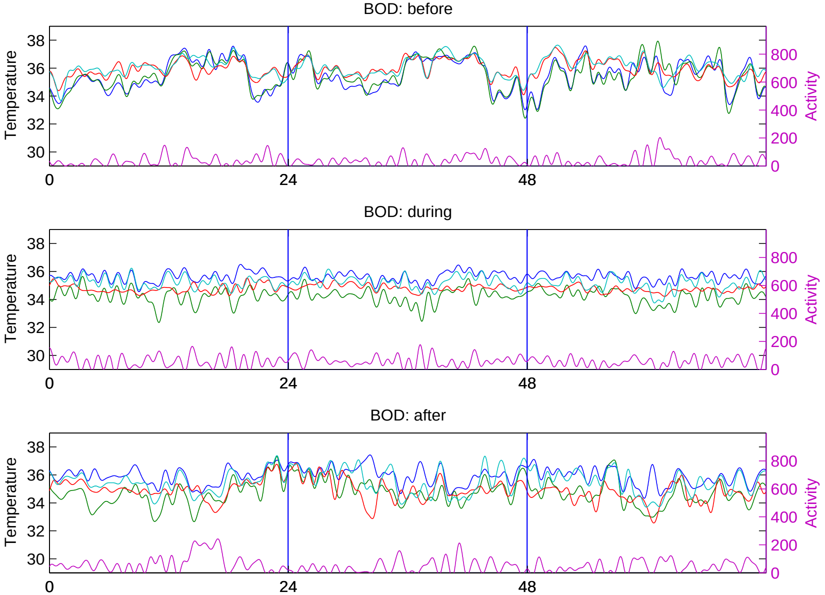


bands 4-6

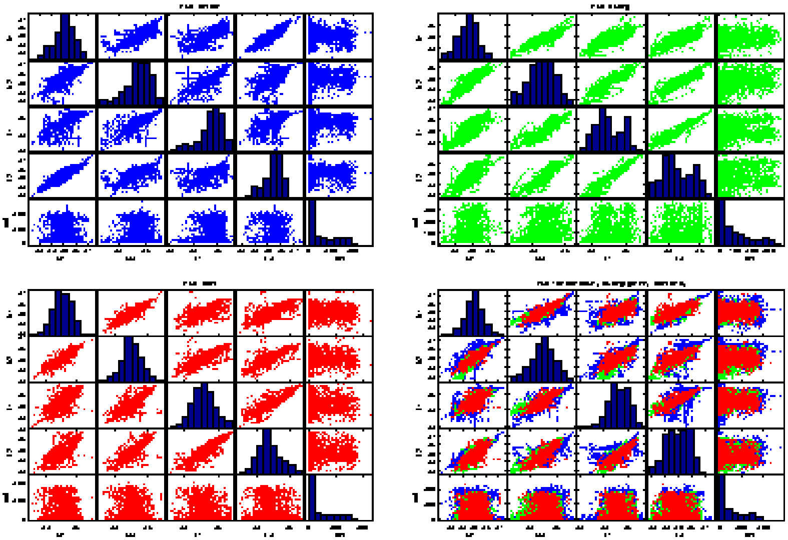


bands 7-9

Temperature and activity Time series before, during and some treatment



Temperatures, before, during and after changes



Simple questions for 1D time series

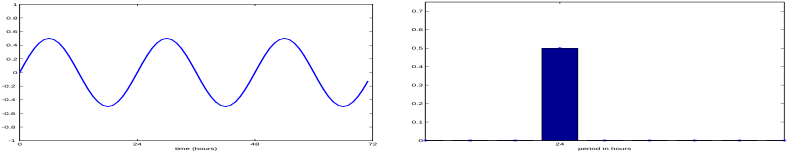
- Main question: Has something changed during and some medical action?
- In this study: effects of circadian cycle on cancer cells.

1. Is there any periodic components in these signals? Yes/No (Detection)? Confidence ?
2. If Yes, How many?
3. What are those components (Periods p_i or Frequencies ν_i , Amplitudes a_i)?

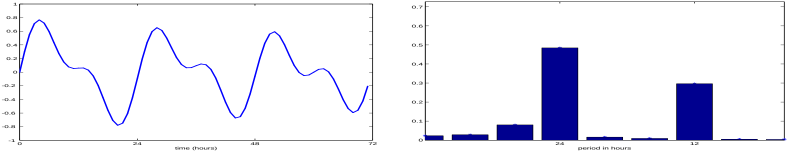
- When questions 1 and 2 are answered, the problem becomes easier: Parameter estimation
- Trying to answer all the three questions at the same time: semi- or Non-Parametric modelling
- Biologists always need uncertainties $\longrightarrow$ Bayesian inference

Simple Analysis tools may not be successful even in very simple cases

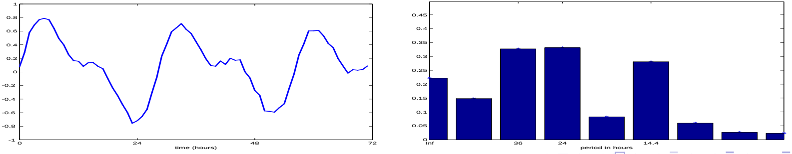
Case of 1 sinusoid



Case of 2 sinusoids+noise



Case of 3 sinusoids+noise



Classical methods: Spectral estimation $S(\omega)$?

- Fast Fourier Transform (FFT):
  $g(t) \longrightarrow FFT \longrightarrow f(\omega) \longrightarrow S(\omega) = |f(\omega)|^2$
  - Advantages: Well-known and understood, fast
  - Drawbacks: linear in frequencies ν,
    but not equidistance in periods
    $\nu = [0, \cdots, N-1] \longrightarrow p = [\infty, 1, \cdots, 1/(N-1)]$
- Autocorrelation function: $\gamma(\tau)$
  - If $g(t)$ is periodic, then $\gamma(\tau)$ is also periodic,
    but much smoother
  - $\gamma(0) = 1 \;\gamma(\tau) \leq \gamma(0), \forall \tau$
- Power spectral density: $\gamma(\tau) \longrightarrow \text{FFT} \longrightarrow S(\omega)$
- Autoregressive (AR), Moving Average (MA) and ARMA models
- Non-stationary GARCH models
- Sum of sinusoidal components

Parametric, Semi- and Non-Parametric models

- Parametric:

$$g(t)=\sum_{k=1}^{K} a_k \sin(2\pi\nu_k t+\phi_k)+\epsilon(t), \quad \boldsymbol{\theta}=\{a_k,\phi_k,\nu_k\}$$

$$g(t)=\sum_{k=1}^{K} a_k \cos(2\pi\nu_k t)+b_k \sin(2\pi\nu_k t)+\epsilon(t), \ \boldsymbol{\theta}=\{a_k,b_k,\nu_k\}$$

$$g(t)=\sum_{k=1}^{K} c_k \exp\left[j2\pi\nu_k t\right]+\epsilon(t), \quad \boldsymbol{\theta}=\{c_k,\nu_k\}, t=0,\cdots,T$$

- Semi-Parametric: $\nu_k=k\nu_0, \nu_0=1/T, K=T \longrightarrow$ DFT
- Non-Parametric: ν_k fixed in a given interval with given precision, so K is fixed but can be as large as necessary.

$$g(t)=\sum_{k=1}^{K} c_k \exp\left[j2\pi\nu_k t\right]+\epsilon(t), \ \boldsymbol{\theta}=\{c_k\} \text{ Linear model}$$

Can we propose a unifying approach for all these problems?

My answer is Yes:

- Identify what your are looking for. (red color $\boldsymbol{f}$)
- Identify what are the data : (blue color $\boldsymbol{g}$)
- Consider the errors (modeling and measurement $\boldsymbol{\epsilon}$)
- Write the Forward model relating them: $\boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\epsilon}$
- Write the expression of the likelihood $p(\boldsymbol{g}|\boldsymbol{f})$
- Translate your prior knowledge on the unknowns in $p(\boldsymbol{f})$
- Use the Bayes rule:

$$p(\boldsymbol{f}|\boldsymbol{g}) = \frac{p(\boldsymbol{g}|\boldsymbol{f})\, p(\boldsymbol{f})}{p(\boldsymbol{g})} \propto p(\boldsymbol{g}|\boldsymbol{f})\, p(\boldsymbol{f})$$

- Infer on $\boldsymbol{f}$ using the posterior $p(\boldsymbol{f}|\boldsymbol{g})$:
  - Maximum A Posteriori (MAP): $\widehat{\boldsymbol{f}} = \arg\max_{\boldsymbol{f}} \{p(\boldsymbol{f}|\boldsymbol{g})\}$
  - Posterior Mean (PM): $\widehat{\boldsymbol{f}} = \int \boldsymbol{f}\, p(\boldsymbol{f}|\boldsymbol{g})\, \mathrm{d}\boldsymbol{f}$

Estimating Periodic Components: Inverse Problems Approach

$$g(t) = \sum_{k=1}^{K} c_k \exp\left[j2\pi\nu_k t\right] + \epsilon(t), \; \boldsymbol{\theta} = \{c_k\} \text{ Linear model}$$

Slight changes of notations: use of periods p_n in place of frequencies ν_k and f_n in place of c_k:

$$g(t) = \sum_{n=1}^{N} f_n \exp\left[j2\pi/p_m t\right] + \epsilon(t), \; t = m\Delta t, m = 1, \cdots, M$$

Defining the vectors: $\boldsymbol{g} = [g_1, \cdots, g_M]'$, $\boldsymbol{\epsilon} = [\epsilon_1, \cdots, \epsilon_M]'$, $\boldsymbol{f} = [f_1, \cdots, f_N]'$ and the matrix $\boldsymbol{H}$: $H_{m,n} = \exp\left[j2\pi/p_m m\Delta t\right]$, we obtain:

$$\boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\epsilon}$$

The objective is to infer on $\boldsymbol{f}$.

Inverse Problems Approach

$$\boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\epsilon}$$

Bayesian approach:

- Assign the Likelihood : $p(\boldsymbol{g}|\boldsymbol{f})$
- Assign the prior law: $p(\boldsymbol{f})$
- Use the Bayes rule : $p(\boldsymbol{f}|\boldsymbol{g}) \propto p(\boldsymbol{g}|\boldsymbol{f})\, p(\boldsymbol{f})$
- MAP:

$$\widehat{\boldsymbol{f}} = \arg\max_{\boldsymbol{f}} \{p(\boldsymbol{f}|\boldsymbol{g})\} = \arg\min_{\boldsymbol{f}} \{J(\boldsymbol{f})\}$$

- Assuming Gaussian noise and Gaussian prior

$$p(\boldsymbol{f}|\boldsymbol{g}) = \mathcal{N}(\boldsymbol{f}|\widehat{\boldsymbol{f}}, \widehat{\boldsymbol{\Sigma}}) \text{ with } \widehat{\boldsymbol{\Sigma}} = (\boldsymbol{H}'\boldsymbol{H} + \lambda\boldsymbol{I})^{-1} \text{ and } \widehat{\boldsymbol{f}} = \arg\max_{\boldsymbol{f}} \{J(\boldsymbol{f})\}$$

$$J(\boldsymbol{f}) = \|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{f}\|^2 + \lambda\|\boldsymbol{f}\|^2$$

- Other priors (Generalized Gaussian, Student-t or Cauchy)
  $J(\boldsymbol{f}) = \|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{f}\|^2 + \lambda\Omega(\boldsymbol{f})$

Bayesian estimation with priors enforcing sparsity

- Sparsity: For any periodic signal, the spectrum is a set of Diracs
- Biological signals related to clock genes: a few independent oscillators
- Spectrum has a few non zero elements in any given interval

$$g(m\Delta t) = \sum_{n=1}^{N} f_n \exp\left[-j2\pi/p_m\, m\Delta t\right] + \epsilon(t), \; m = 1, \cdots, M$$

$\boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\epsilon}$ with $\boldsymbol{f}$ sparse

- The question is now: How to translate sparsity?
- Two solutions: L1 regularization and Bayesian sparsity enforcing priors.
- Three main options in Bayesian: Genealized Gaussian, Student-t, mixtures models

Bayesian estimation with priors enforcing sparsity

- $\boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\epsilon}$ with $\boldsymbol{f}$ sparse
- To translate this information use the heavy tailed prior law Student-t with its hierarchical structure and hidden variables

$$\mathcal{S}t(f_j|\nu) \propto \exp\left[-\frac{\nu+1}{2}\log\left(1+f_j^2/\nu\right)\right]$$

- Infinite Gaussian Scaled Mixture (IGSM) property:

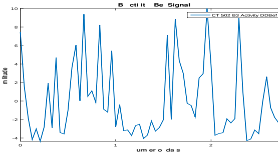
$$\mathcal{S}t(f_j|\nu) \propto = \int_0^\infty \mathcal{N}(f_j|,0,1/z_j)\,\mathcal{G}(z_j|\alpha,\beta)\,\mathrm{d}z_j, \quad \text{with } \alpha=\beta=\nu/2$$

- Hiearchical prior model:

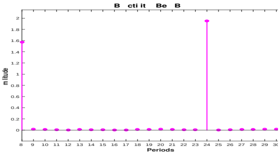
$$p(f_j|z_j) = \mathcal{N}(f_j|0,1/z_j), \quad p(z_j) = \mathcal{G}(z_j|\alpha,\beta)$$

$$\left\{\begin{array}{ll} p(\boldsymbol{f}|\boldsymbol{z}) & = \prod_j p(f_j|z_j) \\ p(\boldsymbol{z}) & = \prod_j p(z_j) \longrightarrow \\ p(\boldsymbol{g}|\boldsymbol{f}) & = \mathcal{N}(\boldsymbol{g}|\boldsymbol{H}\boldsymbol{f},\sigma_\epsilon^2\boldsymbol{I}) \end{array}\right. \qquad p(\boldsymbol{f},\boldsymbol{z}|\boldsymbol{g}) \propto p(\boldsymbol{g}|\boldsymbol{f})p(\boldsymbol{f}|\boldsymbol{z})p(\boldsymbol{z})$$

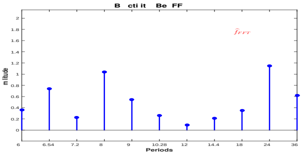
Results on simulated and real activity data



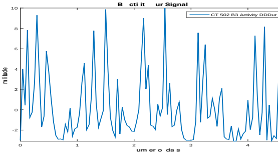
Data BEFORE



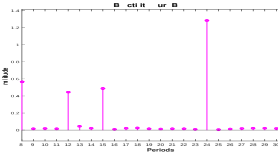
Proposed method



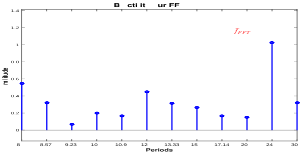
FFT



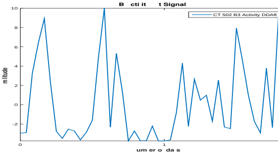
Data DURING



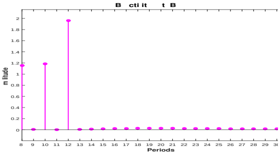
Proposed method



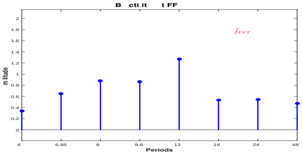
FFT



Data AFTER



Proposed method



FFT

Computed Tomography as Linear Inverse Problem

$\boldsymbol{f}$	all the pixels of the object put together in a vector;
$\boldsymbol{g}$	all the projection data put together in a vector;
$\boldsymbol{g}_k$	the projection at angle ϕ_k, so that $\boldsymbol{g} = [\boldsymbol{g}_1; \boldsymbol{g}_2; \cdots, \boldsymbol{g}_K]$;
$\boldsymbol{H}$	the forward problem matrix: $H_{i,j}$ is the length of the ray i in the pixel j;
$\boldsymbol{H}_k$	the forward problem matrix for the projection at angle ϕ_k, so that $\boldsymbol{H} = [\boldsymbol{H}_1; \boldsymbol{H}_2; \cdots, \boldsymbol{H}_K]$;
$\boldsymbol{\epsilon}$	the noise over all the projection data;
$\boldsymbol{\epsilon}_k$	the noise over the projection at angle ϕ_k so that $\boldsymbol{\epsilon} = [\epsilon_1; \epsilon_2; \cdots, \epsilon_K]$.

$$\boldsymbol{g}_k = \boldsymbol{H}_k \boldsymbol{f} + \boldsymbol{\epsilon}_k \longrightarrow \boldsymbol{g} = \boldsymbol{H} \boldsymbol{f} + \boldsymbol{\epsilon}$$

with $\boldsymbol{H}^t = [\boldsymbol{H}_1^t, \cdots, \boldsymbol{H}_K^t]$, $\boldsymbol{g}^t = [\boldsymbol{g}_1^t, \cdots, \boldsymbol{g}_K^t]$, $\boldsymbol{\epsilon}^t = [\boldsymbol{\epsilon}_1^t, \cdots, \boldsymbol{\epsilon}_K^t]$, where K is the number of projections at angles $\boldsymbol{\phi} = [\phi_1, \cdots, \phi_K]$.

Computed Tomography: BP and FBP

Back-Projection (BP)

$$\widehat{\boldsymbol{f}} = \boldsymbol{H}^t \boldsymbol{g} = \sum_{k=1}^{K} \boldsymbol{H}_k^t \boldsymbol{g}_k.$$

Filtered Back-Projection (FBP)

$$\widehat{\boldsymbol{f}} = \boldsymbol{H}^t \boldsymbol{g} = \sum_{k=1}^{K} \boldsymbol{H}_k^t \widetilde{\boldsymbol{g}}_k.$$

where $\widetilde{\boldsymbol{g}}_k$ is the $|\Omega|$ filtered of $\boldsymbol{g}_k$.
It can also be regarded as the minimum norm solution of $\boldsymbol{H}\boldsymbol{f} = \boldsymbol{g}$:

$$\text{minimize} \quad \|\widehat{\boldsymbol{f}}\|_2^2 \quad \text{subject to} \quad \boldsymbol{H}\boldsymbol{f} = \boldsymbol{g}$$

$$\widehat{\boldsymbol{f}} = \boldsymbol{H}^t (\boldsymbol{H}\boldsymbol{H}^t)^{-1} \boldsymbol{g} = \sum_{k=1}^{K} \boldsymbol{H}_k^t (\boldsymbol{H}_k \boldsymbol{H}_k^t)^{-1} \boldsymbol{g}_k.$$

Computed Tomography: LS and Regularization

Least-Square (LS) and Quadratic Regularization (QR) methods

$$\widehat{\boldsymbol{f}} = \arg\min_{\boldsymbol{f}} \{J(\boldsymbol{f})\}$$

with

$$J(\boldsymbol{f}) = \|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{f}\|_2^2 + \lambda\|\boldsymbol{D}\boldsymbol{f}\|_2^2$$

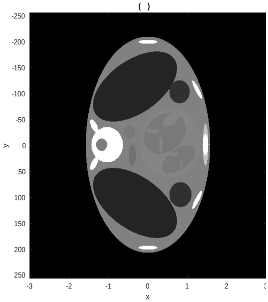
which is given by

$$\widehat{\boldsymbol{f}} = (\boldsymbol{H}^t\boldsymbol{H} + \lambda\boldsymbol{D}^t\boldsymbol{D})^{-1}\boldsymbol{H}^t\boldsymbol{g}.$$

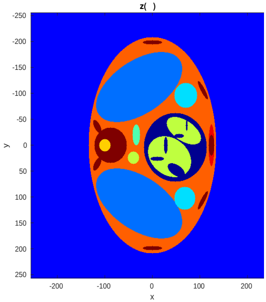
$$J(\boldsymbol{f}) = \sum_{k=1}^{K} \|\boldsymbol{g}_k - \boldsymbol{H}_k\boldsymbol{f}\|_2^2 + \lambda\|\boldsymbol{D}\boldsymbol{f}\|_2^2$$

$$\widehat{\boldsymbol{f}} = \sum_{k=1}^{K} (\boldsymbol{H}_k^t\boldsymbol{H}_k + \lambda\boldsymbol{D}^t\boldsymbol{D})^{-1}\boldsymbol{H}_k^t\boldsymbol{g}.$$

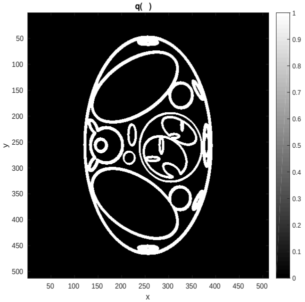
Computed Tomography: Images, Contours, Regions



$f(x, y)$



$z(x, y)$



$q(x, y)$

Next week: Markovian and Hierarchical models