Inverse Problems in Imaging and Computer Vision: From Deterministic Regularization to Probabilistic Bayesian Approaches

Ali Mohammad-Djafari

Groupe Problèmes Inverses
Laboratoire des Signaux et Systèmes
UMR 8506 CNRS - SUPELEC - Univ Paris Sud 11
Supélec, Plateau de Moulon, 91192 Gif-sur-Yvette, FRANCE.

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

WORLDCOM 09, July 13-16, Las Vegas, Nevada, USA
The 2009 World Congres in Computer Science, Computer Engineering and Applied Computing

Content

- Invers problems : Examples and general formulation
- Inversion methods :
  analytical, parametric and non parametric
- Determinitic methods : (Data matching, LS, Regularization)
- Probabilistic methods (Probability matching, Maximum likelihood, Bayesian)
- Bayesian inference approach
- Prior moedels for images
- Bayesian computation
- Applications (Computed Tomography, Image separation)
- Conclusions
- Questions and Discussion

Inverse problems : 3 main examples

- Example 1 :
  Measuring variation of temperature with a therometer
  - $f(t)$ variation of temperature over time
  - $g(t)$ variation of legth of the liquid in thermometer
- Example 2 :
  Making an image with a camera, a microscope or a telescope
  - $f(x,y)$ real scene
  - $g(x,y)$ observed image
- Example 3 : Making an image of the interior of a body
  - $f(x,y)$ a section of a real 3D body $f(x,y,z)$
  - $g_\phi(r)$ a line of observed radiographe $g_\phi(r,z)$

- Example 1 : Deconvolution
- Example 2 : Image restoration
- Example 3 : Image reconstruction

Measuring variation of temperature with a therometer

- $f(t)$ variation of temperature over time
- $g(t)$ variation of legth of the liquid in thermometer
- Forward model : Convolution

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

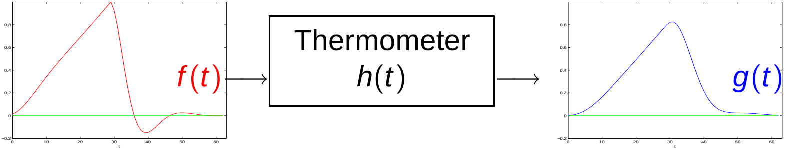
  $h(t)$: impulse response of the measurement system
- Inverse problem : Deconvolution

  Given the forward model $\mathcal{H}$ (impulse response $h(t)$))
  and a set of data $g(t_i), i = 1, \cdots, M$
  find $f(t)$

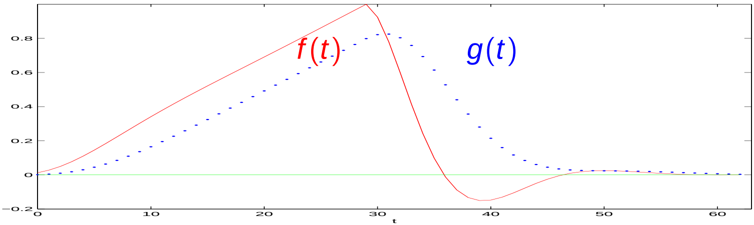
Measuring variation of temperature with a therometer

Forward model : Convolution

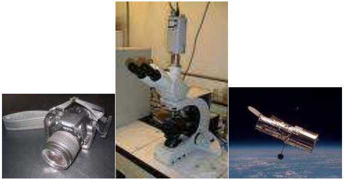
$$g(t) = \int f(t')\, h(t-t')\, \mathrm{d}t' + \epsilon(t)$$



Inversion : Deconvolution



Making an image with a camera, a microscope or a telescope



- $f(x,y)$ real scene
- $g(x,y)$ observed image
- Forward model : Convolution

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

  $h(x,y)$: Point Spread Function (PSF) of the imaging system

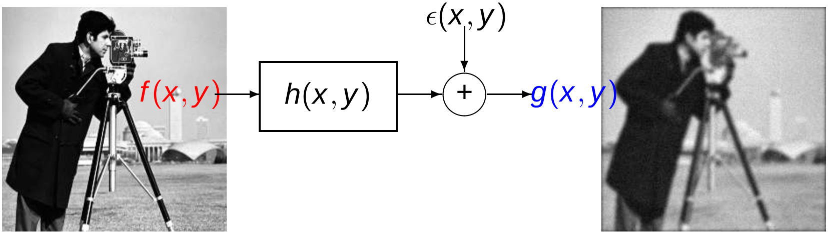
- Inverse problem : Image restoration

  Given the forward model $\mathcal{H}$ (PSF $h(x,y)$))
  and a set of data $g(x_i,y_i), i = 1,\cdots,M$
  find $f(x,y)$

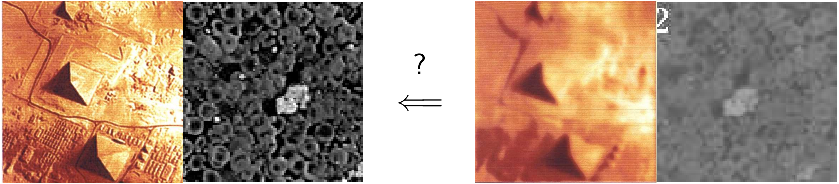
Making an image with an unfocused camera

Forward model : 2D Convolution

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

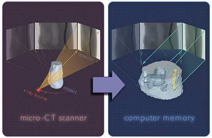


Inversion : Deconvolution



Making an image of the interior of a body

- $f(x,y)$ a section of a real 3D body $f(x,y,z)$
- $g_\phi(r)$ a line of observed radiographe $g_\phi(r,z)$



- Forward model :
  Line integrals or Radon Transform

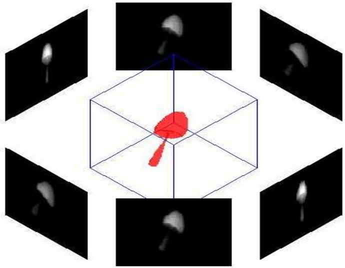
$$
\begin{aligned}
g_\phi(r) &= \int_{L_{r,\phi}} f(x,y)\, \mathrm{d}l + \epsilon_\phi(r) \\
&= \iint f(x,y)\, \delta(r - x\cos\phi - y\sin\phi)\, \mathrm{d}x\, \mathrm{d}y + \epsilon_\phi(r)
\end{aligned}
$$

- Inverse problem : Image reconstruction

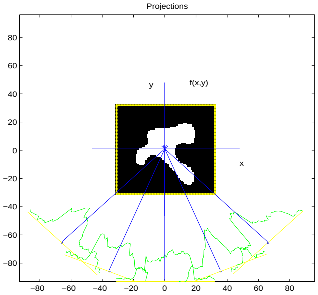
  Given the forward model $\mathcal{H}$ (Radon Transform) and
  a set of data $g_{\phi_i}(r), i = 1, \cdots, M$
  find $f(x,y)$

2D and 3D Computed Tomography

3D



2D



$$g_\phi(r_1,r_2) = \int_{\mathcal{L}_{r_1,r_2,\phi}} f(x,y,z)\,\mathrm{d}l \qquad g_\phi(r) = \int_{\mathcal{L}_{r,\phi}} f(x,y)\,\mathrm{d}l$$

Forward probelm : $f(x,y)$ or $f(x,y,z) \longrightarrow g_\phi(r)$ or $g_\phi(r_1,r_2)$

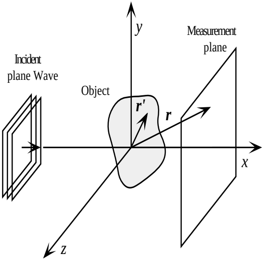
Inverse problem : $g_\phi(r)$ or $g_\phi(r_1,r_2) \longrightarrow f(x,y)$ or $f(x,y,z)$

Microwave or ultrasound imaging

Mesaurs : diffracted wave by the object $\phi_d(\boldsymbol{r}_i)$

Unknown quantity : $f(\boldsymbol{r}) = k_0^2(n^2(\boldsymbol{r}) - 1)$

Intermediate quantity : $\phi(\boldsymbol{r})$

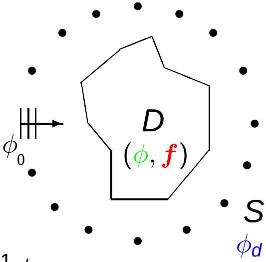


$$\phi_d(\boldsymbol{r}_i) = \iint_D G_m(\boldsymbol{r}_i, \boldsymbol{r}')\phi(\boldsymbol{r}') f(\boldsymbol{r}') \,\mathrm{d}\boldsymbol{r}', \; \boldsymbol{r}_i \in S$$

$$\phi(\boldsymbol{r}) = \phi_0(\boldsymbol{r}) + \iint_D G_o(\boldsymbol{r}, \boldsymbol{r}')\phi(\boldsymbol{r}') f(\boldsymbol{r}') \,\mathrm{d}\boldsymbol{r}', \; \boldsymbol{r} \in D$$

Born approximation $(\phi(\boldsymbol{r}') \simeq \phi_0(\boldsymbol{r}'))$ **:**

$$\phi_d(\boldsymbol{r}_i) = \iint_D G_m(\boldsymbol{r}_i, \boldsymbol{r}')\phi_0(\boldsymbol{r}') f(\boldsymbol{r}') \,\mathrm{d}\boldsymbol{r}', \; \boldsymbol{r}_i \in S$$



Discretization :

$$\left\{\begin{array}{l} \phi_d = \boldsymbol{G}_m \boldsymbol{F} \boldsymbol{\phi} \\ \boldsymbol{\phi} = \phi_0 + \boldsymbol{G}_o \boldsymbol{F} \boldsymbol{\phi} \end{array}\right. \longrightarrow \left\{\begin{array}{l} \phi_d = \boldsymbol{H}(\boldsymbol{f}) \\ \text{with } \boldsymbol{F} = \mathrm{diag}(\boldsymbol{f}) \\ \boldsymbol{H}(\boldsymbol{f}) = \boldsymbol{G}_m \boldsymbol{F} (\boldsymbol{I} - \boldsymbol{G}_o \boldsymbol{F})^{-1} \phi_0 \end{array}\right.$$

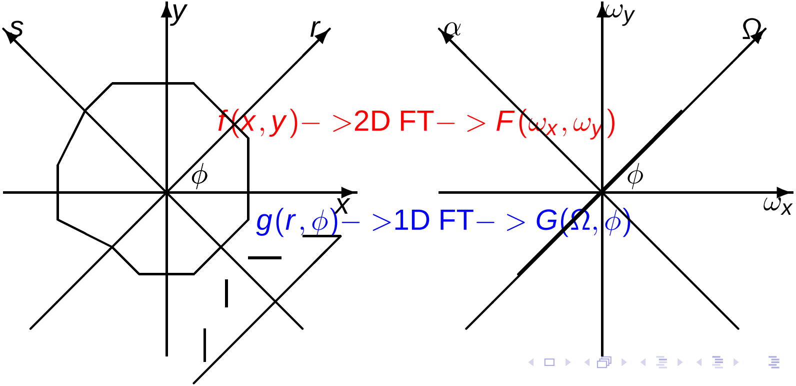
Fourier Synthesis in X ray Tomography

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

$$G(\Omega,\phi) = \int g(r,\phi)\exp\{-j\Omega r\}\,\mathrm{d}r$$

$$F(\omega_x,\omega_y) = \iint f(x,y)\exp\{-j\omega_x x,\omega_y y\}\,\mathrm{d}x\,\mathrm{d}y$$

$$F(\omega_x,\omega_y) = P(\Omega,\phi) \quad \text{for} \quad \omega_x = \Omega\cos\phi \quad \text{and} \quad \omega_y = \Omega\sin\phi$$

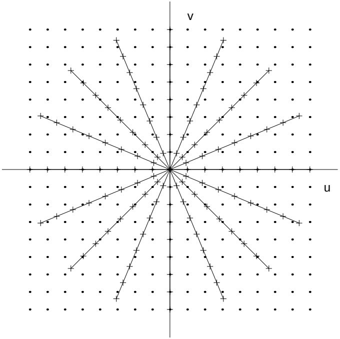


$f(x,y) -> \text{2D FT} -> F(\omega_x,\omega_y)$

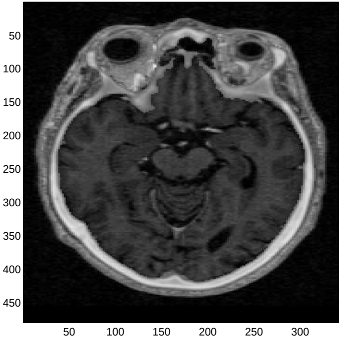
$g(r,\phi) -> \text{1D FT} -> G(\Omega,\phi)$

Fourier Synthesis in X ray tomography

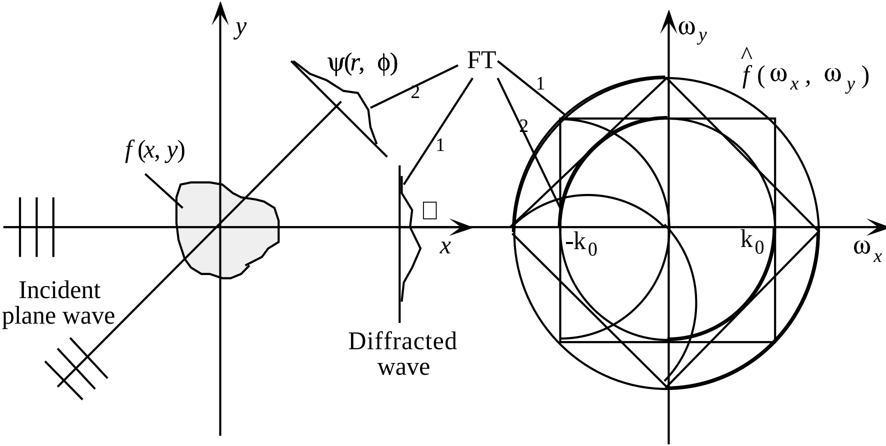
$$F(\omega_x, \omega_y) = \iint f(x,y) \exp\left\{-j\omega_x x, \omega_y y\right\} \, dx \, dy$$



$$? \\ \Longrightarrow$$

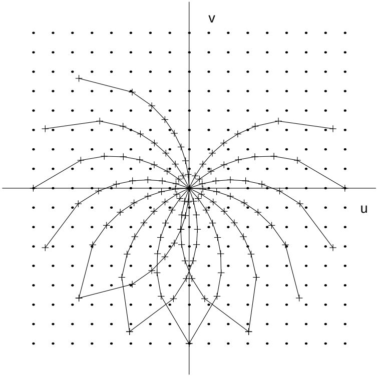


Fourier Synthesis in Diffraction tomography



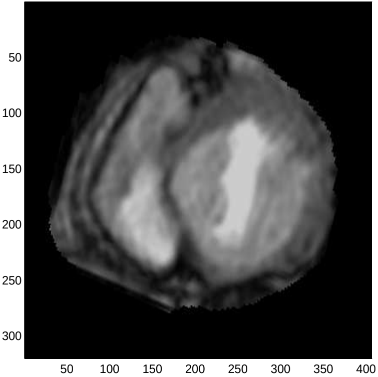
Fourier Synthesis in Diffraction tomography

$$F(\omega_x, \omega_y) = \iint f(x, y) \exp\left\{-j\omega_x x, \omega_y y\right\} \, dx \, dy$$



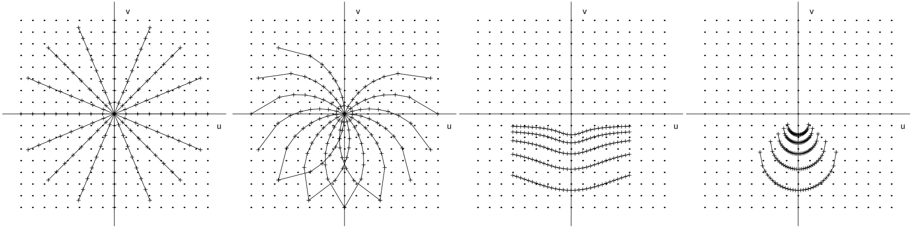
?

$\Longrightarrow$



Fourier Synthesis in different imaging systems

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



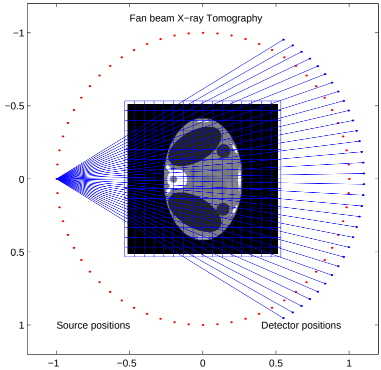
X ray Tomography Diffraction Eddy current SAR & Radar

Invers Problems : other examples and applications

- X ray, Gamma ray Computed Tomography (CT)
- Microwave and ultrasound tomography
- Positron emission tomography (PET)
- Magnetic resonance imaging (MRI)
- Photoacoustic imaging
- Radio astronomy
- Geophysical imaging
- Non Destructive Evaluation (NDE) and Testing (NDT) techniques in industry
- Hyperspectral imaging
- Earth observation methods (Radar, SAR, IR, ...)
- Survey and tracking in security systems

Computed tomography (CT)

A Multislice CT Scanner

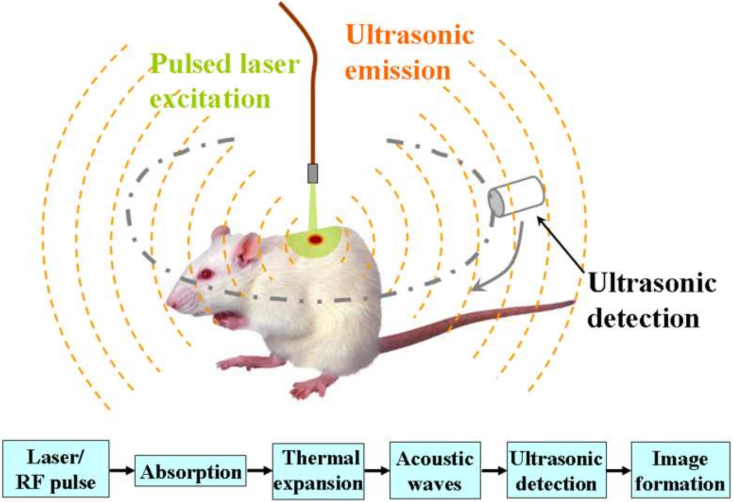


$$g(s_i) = \int_{L_i} f(\boldsymbol{r}) \; \mathrm{d}l_i + \epsilon(s_i)$$

Discretization

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

Photoacoustic imaging



Positron emission tomography (PET)

Magnetic resonance imaging (MRI)

Nuclear magnetic resonance imaging (NMRI), Para-sagittal MRI of the head

Radio astronomy (interferometry imaging systems)

The Very Large Array in New Mexico, an example of a radio telescope.

General formulation of inverse problems

- General non linear inverse problems :

$$g(\boldsymbol{s}) = [\mathcal{H} f(\boldsymbol{r})](\boldsymbol{s}) + \epsilon(\boldsymbol{s}), \quad \boldsymbol{r} \in \mathcal{R}, \quad \boldsymbol{s} \in \mathcal{S}$$

- Linear models :

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

  If $h(\boldsymbol{r}, \boldsymbol{s}) = h(\boldsymbol{r} - \boldsymbol{s}) \longrightarrow$ Convolution.

- Discrete data :

$$g(\boldsymbol{s}_i) = \int h(\boldsymbol{s}_i, \boldsymbol{r})\, f(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r} + \epsilon(\boldsymbol{s}_i), \quad i = 1, \cdots, m$$

- Inversion : Given the forward model $\mathcal{H}$ and the data $\boldsymbol{g} = \{g(\boldsymbol{s}_i), i = 1, \cdots, m)\}$ estimate $f(\boldsymbol{r})$
- Well-posed and Ill-posed problems (Hadamard) : existance, uniqueness and stability
- Need for prior information

Analytical methods (mathematical physics)

$$g(\boldsymbol{s}_i) = \int h(\boldsymbol{s}_i, \boldsymbol{r})\, f(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r} + \epsilon(\boldsymbol{s}_i), \quad i = 1, \cdots, m$$

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

$$\widehat{f}(\boldsymbol{r}) = \int w(\boldsymbol{s}, \boldsymbol{r})\, g(\boldsymbol{s})\, \mathrm{d}\boldsymbol{s}$$

$w(\boldsymbol{s}, \boldsymbol{r})$ minimizing a criterion :

$$\begin{aligned}
Q(w(\boldsymbol{s}, \boldsymbol{r})) &= \left\| g(\boldsymbol{s}) - [\mathcal{H}\widehat{f}(\boldsymbol{r})](\boldsymbol{s}) \right\|_2^2 = \int \left| g(\boldsymbol{s}) - [\mathcal{H}\widehat{f}(\boldsymbol{r})](\boldsymbol{s}) \right|^2 \mathrm{d}\boldsymbol{s} \\
&= \int \left| g(\boldsymbol{s}) - \int h(\boldsymbol{s}, \boldsymbol{r})\, \widehat{f}(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r} \right|^2 \mathrm{d}\boldsymbol{s} \\
&= \int \left| g(\boldsymbol{s}) - \int h(\boldsymbol{s}, \boldsymbol{r}) \left[\int w(\boldsymbol{s}, \boldsymbol{r})\, g(\boldsymbol{s})\, \mathrm{d}\boldsymbol{s} \right] \mathrm{d}\boldsymbol{r} \right|^2 \mathrm{d}\boldsymbol{s} \\
&= \int \left| g(\boldsymbol{s}) - \int \int h(\boldsymbol{s}, \boldsymbol{r}) w(\boldsymbol{s}, \boldsymbol{r})\, g(\boldsymbol{s})\, \mathrm{d}\boldsymbol{s}\, \mathrm{d}\boldsymbol{r} \right|^2 \mathrm{d}\boldsymbol{s}
\end{aligned}$$

Analytical methods

- Trivial solution :

$$w(\boldsymbol{s},\boldsymbol{r}) = h^{-1}(\boldsymbol{s},\boldsymbol{r})$$

  Example : Fourier Transform :

$$g(\boldsymbol{s}) = \int f(\boldsymbol{r}) \exp\{-j\boldsymbol{s}.\boldsymbol{r}\} \,\mathrm{d}\boldsymbol{r}$$

$$h(\boldsymbol{s},\boldsymbol{r}) = \exp\{-j\boldsymbol{s}.\boldsymbol{r}\} \longrightarrow w(\boldsymbol{s},\boldsymbol{r}) = \exp\{+j\boldsymbol{s}.\boldsymbol{r}\}$$

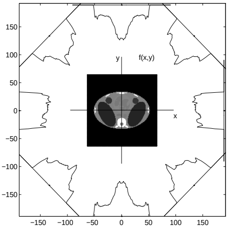
$$\hat{f}(\boldsymbol{r}) = \int g(\boldsymbol{s}) \exp\{+j\boldsymbol{s}.\boldsymbol{r}\} \,\mathrm{d}\boldsymbol{s}$$

- Known classical solutions for specific expressions of $h(\boldsymbol{s},\boldsymbol{r})$:
  - 1D cases : 1D Fourier, Hilbert, Weil, Melin, ...
  - 2D cases : 2D Fourier, Radon, ...

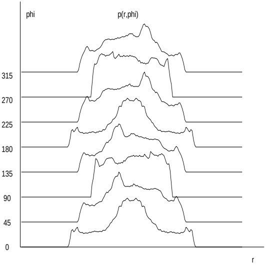
X ray Tomography

$$g(r,\phi) = -\ln\left(\frac{I}{I_0}\right) = \int_{L_{r,\phi}} f(x,y)\ \mathrm{d}l$$

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



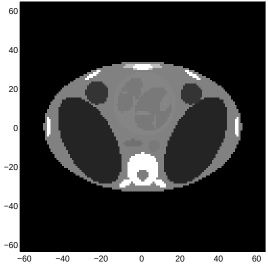
$$f(x,y) \longrightarrow \boxed{\text{RT}} \longrightarrow g(r,\phi)$$



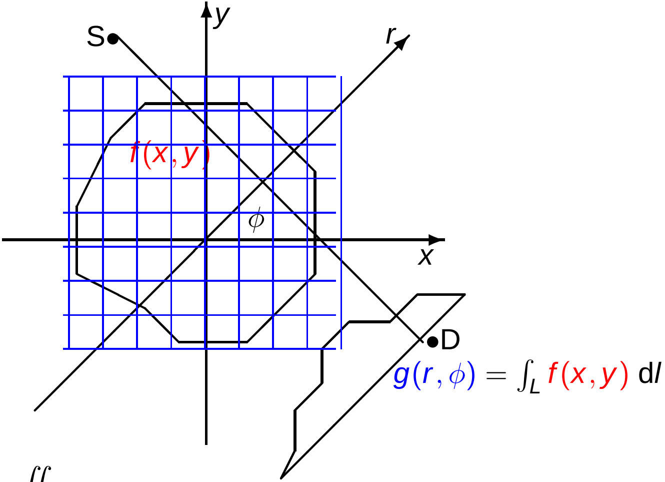
IRT

?

$\Longrightarrow$



Analytical Inversion methods



Radon :

$$
\begin{aligned}
g(r,\phi) &= \iint_D f(x,y)\,\delta(r - x\cos\phi - y\sin\phi)\;dx\;dy \\
f(x,y) &= \left(-\frac{1}{2\pi^2}\right)\int_0^{\pi}\int_{-\infty}^{+\infty}\frac{\frac{\partial}{\partial r}g(r,\phi)}{(r - x\cos\phi - y\sin\phi)}\;dr\;d\phi
\end{aligned}
$$

Filtered Backprojection method

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

Derivation $\mathcal{D}$: $\quad \bar{g}(r,\phi) = \dfrac{\partial g(r,\phi)}{\partial r}$

Hilbert Transform $\mathcal{H}$: $\quad g_1(r',\phi) = \dfrac{1}{\pi} \displaystyle\int_0^{\infty} \frac{\bar{g}(r,\phi)}{(r - r')}\, \mathrm{d}r$

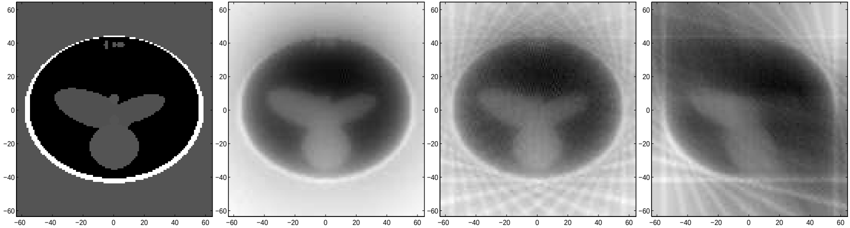
Backprojection $\mathcal{B}$: $\quad f(x,y) = \dfrac{1}{2\pi} \displaystyle\int_0^{\pi} g_1(r' = x\cos\phi + y\sin\phi, \phi)\, \mathrm{d}\phi$

$$f(x,y) = \mathcal{B}\, \mathcal{H}\, \mathcal{D}\, g(r,\phi) = \mathcal{B}\, \mathcal{F}_1^{-1}\, |\Omega|\, \mathcal{F}_1\, g(r,\phi)$$

- Backprojection of filtered projections :

$$\xrightarrow{g(r,\phi)} \boxed{\begin{array}{c}\text{FT}\\ \mathcal{F}_1\end{array}} \longrightarrow \boxed{\begin{array}{c}\text{Filter}\\ |\Omega|\end{array}} \longrightarrow \boxed{\begin{array}{c}\text{IFT}\\ \mathcal{F}_1^{-1}\end{array}} \xrightarrow{g_1(r,\phi)} \boxed{\begin{array}{c}\text{Backprojection}\\ \mathcal{B}\end{array}} \xrightarrow{f(x,y)}$$

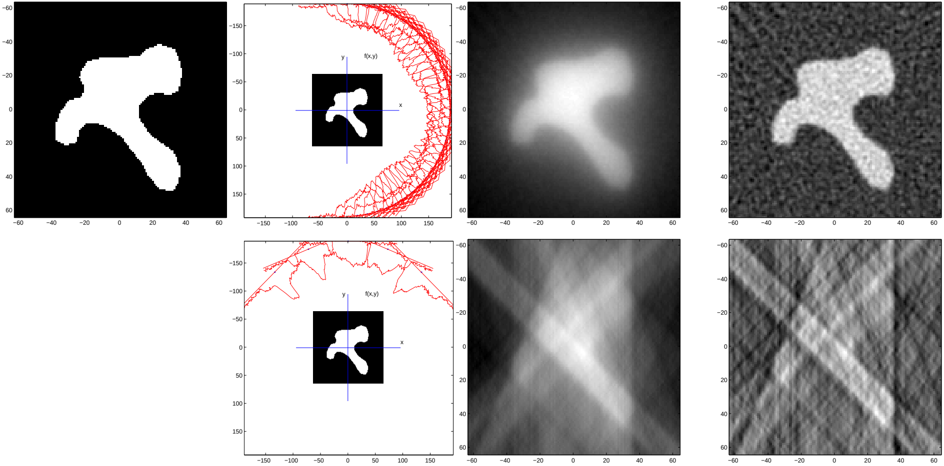
Limitations : Limited angle or noisy data



Original | 64 proj. | 16 proj. | 8 proj. $[0, \pi/2]$

- Limited angle or noisy data
- Accounting for detector size
- Other measurement geometries : fan beam, ...

Limitations : Limited angle or noisy data



Original Data Backprojection Filtered Backprojectio

Parametric methods

- $f(\boldsymbol{r})$ is described in a parametric form with a very few number of parameters $\boldsymbol{\theta}$ and one searches $\widehat{\boldsymbol{\theta}}$ which minimizes a criterion such as :
- Least Squares (LS) : $Q(\boldsymbol{\theta}) = \sum_i |g_i - [\mathcal{H} f(\boldsymbol{\theta})]_i|^2$
- Robust criteria : $Q(\boldsymbol{\theta}) = \sum_i \phi\left(|g_i - [\mathcal{H} f(\boldsymbol{\theta})]_i|\right)$
  with different functions ϕ (L_1, Hubert, ...).
- Likelihood : $\mathcal{L}(\boldsymbol{\theta}) = -\ln p(\boldsymbol{g}|\boldsymbol{\theta})$
- Penalized likelihood : $\mathcal{L}(\boldsymbol{\theta}) = -\ln p(\boldsymbol{g}|\boldsymbol{\theta}) + \lambda\Omega(\boldsymbol{\theta})$

Examples :

- Spectrometry : $f(t)$ modelled as a sum og gaussians
  $f(t) = \sum_{k=1}^{K} a_k \mathcal{N}(t|\mu_k, v_k)$ $\quad \boldsymbol{\theta} = \{a_k, \mu_k, v_k\}$
- Tomography in CND : $f(x,y)$ is modelled as a superposition of circular or elleiptical discs
  $\boldsymbol{\theta} = \{a_k, \mu_k, r_k\}$

Non parametric methods

$$g(\boldsymbol{s}_i) = \int h(\boldsymbol{s}_i, \boldsymbol{r})\, f(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r} + \epsilon(\boldsymbol{s}_i), \quad i = 1, \cdots, M$$

- $f(\boldsymbol{r})$ is assumed to be well approximated by

$$f(\boldsymbol{r}) \simeq \sum_{j=1}^{N} f_j\, b_j(\boldsymbol{r})$$

with $\{b_j(\boldsymbol{r})\}$ a basis or any other set of known functions

$$g(\boldsymbol{s}_i) = g_i \simeq \sum_{j=1}^{N} f_j \int h(\boldsymbol{s}_i, \boldsymbol{r})\, b_j(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r}, \quad i = 1, \cdots, M$$

$$\boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\epsilon} \text{ with } H_{ij} = \int h(\boldsymbol{s}_i, \boldsymbol{r})\, b_j(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r}$$

- $\boldsymbol{H}$ is huge dimensional
- LS solution : $\widehat{\boldsymbol{f}} = \arg\min_{\boldsymbol{f}} \{Q(\boldsymbol{f})\}$ with
  $Q(\boldsymbol{f}) = \sum_i |g_i - [\boldsymbol{H}\boldsymbol{f}]_i|^2 = \|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{f}\|^2$
  does not give satisfactory result.

Inversion : Deterministic methods
Data matching

- Observation model
  $g_i = h_i(\boldsymbol{f}) + \epsilon_i, \quad i = 1, \ldots, M \longrightarrow \boldsymbol{g} = \boldsymbol{H}(\boldsymbol{f}) + \boldsymbol{\epsilon}$
- Misatch between data and output of the model $\Delta(\boldsymbol{g}, \boldsymbol{H}(\boldsymbol{f}))$

$$\widehat{\boldsymbol{f}} = \arg\min_{\boldsymbol{f}} \{\Delta(\boldsymbol{g}, \boldsymbol{H}(\boldsymbol{f}))\}$$

- Examples :

– LS $\quad \Delta(\boldsymbol{g}, \boldsymbol{H}(\boldsymbol{f})) = \|\boldsymbol{g} - \boldsymbol{H}(\boldsymbol{f})\|^2 = \sum_i |g_i - h_i(\boldsymbol{f})|^2$

– L_p $\quad \Delta(\boldsymbol{g}, \boldsymbol{H}(\boldsymbol{f})) = \|\boldsymbol{g} - \boldsymbol{H}(\boldsymbol{f})\|^p = \sum_i |g_i - h_i(\boldsymbol{f})|^p, \quad 1 < p < 2$

– KL $\quad \Delta(\boldsymbol{g}, \boldsymbol{H}(\boldsymbol{f})) = \sum_i g_i \ln \frac{g_i}{h_i(\boldsymbol{f})}$

- In general, does not give satisfactory results for inverse problems.

Regularization theory

Inverse problems = Ill posed problems

$\longrightarrow$ Need for prior information

Functional space (Tikhonov) :

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

Finite dimensional space (Philips & Towmey) : $\boldsymbol{g} = \boldsymbol{H}(\boldsymbol{f}) + \boldsymbol{\epsilon}$

- Minimum norme LS (MNLS) : $J(\boldsymbol{f}) = ||\boldsymbol{g} - \boldsymbol{H}(\boldsymbol{f})||^2 + \lambda||\boldsymbol{f}||^2$
- Classical regularization : $J(\boldsymbol{f}) = ||\boldsymbol{g} - \boldsymbol{H}(\boldsymbol{f})||^2 + \lambda||\boldsymbol{D}\boldsymbol{f}||^2$
- More general regularization :

$$J(\boldsymbol{f}) = \mathcal{Q}(\boldsymbol{g} - \boldsymbol{H}(\boldsymbol{f})) + \lambda\Omega(\boldsymbol{D}\boldsymbol{f})$$

or

$$J(\boldsymbol{f}) = \Delta_1(\boldsymbol{g}, \boldsymbol{H}(\boldsymbol{f})) + \lambda\Delta_2(\boldsymbol{f}, \boldsymbol{f}_\infty)$$

Limitations :

- Errors are considered implicitly white and Gaussian
- Limited prior information on the solution
- Lack of tools for the determination of the hyperparameters

Inversion : Probabilistic methods

Taking account of errors and uncertainties $\longrightarrow$ Probability theory

- Maximum Likelihood (ML)
- Minimum Inaccuracy (MI)
- Probability Distribution Matching (PDM)
- Maximum Entropy (ME) and Information Theory (IT)
- Bayesian Inference (BAYES)

Advantages :

- Explicit account of the errors and noise
- A large class of priors via explicit or implicit modeling
- A coherent approach to combine information content of the data and priors

Limitations :

- Practical implementation and cost of calculation

Bayesian estimation approach

$$\mathcal{M}: \qquad \boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\epsilon}$$

- Observation model $\mathcal{M}$ + Hypothesis on the noise $\boldsymbol{\epsilon} \longrightarrow$ $p(\boldsymbol{g}|\boldsymbol{f};\mathcal{M}) = p_{\boldsymbol{\epsilon}}(\boldsymbol{g} - \boldsymbol{H}\boldsymbol{f})$
- A priori information $\quad p(\boldsymbol{f}|\mathcal{M})$
- Bayes : $\quad p(\boldsymbol{f}|\boldsymbol{g};\mathcal{M}) = \dfrac{p(\boldsymbol{g}|\boldsymbol{f};\mathcal{M})\, p(\boldsymbol{f}|\mathcal{M})}{p(\boldsymbol{g}|\mathcal{M})}$

Link with regularization :

Maximum A Posteriori (MAP) :

$$\begin{aligned}
\widehat{\boldsymbol{f}} &= \arg\max_{\boldsymbol{f}} \{p(\boldsymbol{f}|\boldsymbol{g})\} = \arg\max_{\boldsymbol{f}} \{p(\boldsymbol{g}|\boldsymbol{f})\, p(\boldsymbol{f})\} \\
&= \arg\min_{\boldsymbol{f}} \{-\ln p(\boldsymbol{g}|\boldsymbol{f}) - \ln p(\boldsymbol{f})\}
\end{aligned}$$

with $\quad Q(\boldsymbol{g}, \boldsymbol{H}\boldsymbol{f}) = -\ln p(\boldsymbol{g}|\boldsymbol{f}) \quad$ and $\quad \lambda\Omega(\boldsymbol{f}) = -\ln p(\boldsymbol{f})$

Case of linear models and Gaussian priors

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

- Hypothesis on the noise : $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \sigma_\epsilon^2 \boldsymbol{I}) \longrightarrow$
  $p(\boldsymbol{g}|\boldsymbol{f}) \propto \exp\left\{-\frac{1}{2\sigma_\epsilon^2}\|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{f}\|^2\right\}$
- Hypothesis on $\boldsymbol{f}$: $\boldsymbol{f} \sim \mathcal{N}(0, \sigma_f^2 (\boldsymbol{D}^t\boldsymbol{D})^{-1}) \longrightarrow$
  $p(\boldsymbol{f}) \propto \exp\left\{-\frac{1}{2\sigma_f^2}\|\boldsymbol{D}\boldsymbol{f}\|^2\right\}$
- A posteriori :
  $$p(\boldsymbol{f}|\boldsymbol{g}) \propto \exp\left\{-\frac{1}{2\sigma_\epsilon^2}\|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{f}\|^2 \frac{1}{2\sigma_f^2}\|\boldsymbol{D}\boldsymbol{f}\|^2\right\}$$
- MAP : $\widehat{\boldsymbol{f}} = \arg\max_{\boldsymbol{f}} \{p(\boldsymbol{f}|\boldsymbol{g})\} = \arg\min_{\boldsymbol{f}} \{J(\boldsymbol{f})\}$
  with $J(\boldsymbol{f}) = \|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{f}\|^2 + \lambda\|\boldsymbol{D}\boldsymbol{f}\|^2, \qquad \lambda = \frac{\sigma_\epsilon^2}{\sigma_f^2}$
- Advantage : characterization of the solution

$$\boldsymbol{f}|\boldsymbol{g} \sim \mathcal{N}(\widehat{\boldsymbol{f}}, \widehat{\boldsymbol{P}}) \text{ with } \widehat{\boldsymbol{f}} = \widehat{\boldsymbol{P}}\boldsymbol{H}^t\boldsymbol{g}, \quad \widehat{\boldsymbol{P}} = \left(\boldsymbol{H}^t\boldsymbol{H} + \lambda\boldsymbol{D}^t\boldsymbol{D}\right)^{-1}$$

MAP estimation with other priors :

$$\widehat{\boldsymbol{f}} = \arg\min_{\boldsymbol{f}} \{J(\boldsymbol{f})\} \quad \text{avec} \quad J(\boldsymbol{f}) = \|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{f}\|^2 + \lambda\Omega(\boldsymbol{f})$$

Separable priors :

- Gaussian : $p(f_j) \propto \exp\left\{-\alpha|f_j|^2\right\} \longrightarrow \Omega(\boldsymbol{f}) = \alpha\sum_j |f_j|^2$
- Gamma : $p(f_j) \propto f_j^\alpha \exp\left\{-\beta f_j\right\} \longrightarrow \Omega(\boldsymbol{f}) = \alpha\sum_j \ln f_j + \beta f_j$
- Beta :
  $p(f_j) \propto f_j^\alpha (1-f_j)^\beta \longrightarrow \Omega(\boldsymbol{f}) = \alpha\sum_j \ln f_j + \beta\sum_j \ln(1-f_j)$
- Generalized Gaussian :
  $p(f_j) \propto \exp\left\{-\alpha|f_j|^p\right\}, \quad 1 < p < 2 \longrightarrow \quad \Omega(\boldsymbol{f}) = \alpha\sum_j |f_j|^p,$

Markovian models :

$$p(f_j|\boldsymbol{f}) \propto \exp\left\{-\alpha\sum_{i\in N_j}\phi(f_j,f_i)\right\} \longrightarrow \quad \Omega(\boldsymbol{f}) = \alpha\sum_j\sum_{i\in N_j}\phi(f_j,f_i),$$

MAP estimation with markovien priors :

$$\widehat{\boldsymbol{f}} = \arg\min_{\boldsymbol{f}} \{J(\boldsymbol{f})\} \quad \text{with} \quad J(\boldsymbol{f}) = \|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{f}\|^2 + \lambda\Omega(\boldsymbol{f})$$

$$\Omega(\boldsymbol{f}) = \sum_j \phi(f_j - f_{j-1})$$

with $\phi(t)$:

Convex functions :

$$|t|^\alpha, \; \sqrt{1+t^2} - 1, \; \log(\cosh(t)), \; \begin{cases} t^2 & |t| \leq T \\ 2T|t| - T^2 & |t| > T \end{cases}$$

or Non convex functions :

$$\log(1+t^2), \quad \frac{t^2}{1+t^2}, \quad \arctan(t^2), \quad \begin{cases} t^2 & |t| \leq T \\ T^2 & |t| > T \end{cases}$$

Main advantages of the Bayesian approach

- MAP = Regularization
- Posterior mean ? Marginal MAP ?
- More information in the posterior law than only its mode or its mean
- Meaning and tools for estimating hyper parameters
- Meaning and tools for model selection
- More specific and specialized priors, particularly through the hidden variables
- More computational tools :
  - Expectation-Maximization for computing the maximum likelihood parameters
  - MCMC for posterior exploration
  - Variational Bayes for analytical computation of the posterior marginals
  - ...

Full Bayesian approach

$$\mathcal{M}: \quad \boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\epsilon}$$

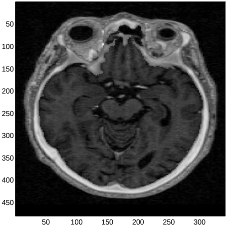
- Forward & errors model : $\longrightarrow p(\boldsymbol{g}|\boldsymbol{f}, \boldsymbol{\theta}_1; \mathcal{M})$
- Prior models $\longrightarrow p(\boldsymbol{f}|\boldsymbol{\theta}_2; \mathcal{M})$
- Hyperparameters $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) \longrightarrow p(\boldsymbol{\theta}|\mathcal{M})$
- Bayes : $\longrightarrow p(\boldsymbol{f}, \boldsymbol{\theta}|\boldsymbol{g}; \mathcal{M}) = \frac{p(\boldsymbol{g}|\boldsymbol{f},\boldsymbol{\theta};\mathcal{M})\, p(\boldsymbol{f}|\boldsymbol{\theta};\mathcal{M})\, p(\boldsymbol{\theta}|\mathcal{M})}{p(\boldsymbol{g}|\mathcal{M})}$
- Joint MAP : $\quad (\widehat{\boldsymbol{f}}, \widehat{\boldsymbol{\theta}}) = \arg\max_{(\boldsymbol{f},\boldsymbol{\theta})} \{p(\boldsymbol{f}, \boldsymbol{\theta}|\boldsymbol{g}; \mathcal{M})\}$
- Marginalization : $\begin{cases} p(\boldsymbol{f}|\boldsymbol{g}; \mathcal{M}) & = \int p(\boldsymbol{f}, \boldsymbol{\theta}|\boldsymbol{g}; \mathcal{M})\, \mathrm{d}\boldsymbol{f} \\ p(\boldsymbol{\theta}|\boldsymbol{g}; \mathcal{M}) & = \int p(\boldsymbol{f}, \boldsymbol{\theta}|\boldsymbol{g}; \mathcal{M})\, \mathrm{d}\boldsymbol{\theta} \end{cases}$
- Posterior means : $\begin{cases} \widehat{\boldsymbol{f}} & = \int \boldsymbol{f}\, p(\boldsymbol{f}, \boldsymbol{\theta}|\boldsymbol{g}; \mathcal{M})\, \mathrm{d}\boldsymbol{f}\, \mathrm{d}\boldsymbol{\theta} \\ \widehat{\boldsymbol{\theta}} & = \int \boldsymbol{\theta}\, p(\boldsymbol{f}, \boldsymbol{\theta}|\boldsymbol{g}; \mathcal{M})\, \mathrm{d}\boldsymbol{f}\, \mathrm{d}\boldsymbol{\theta} \end{cases}$
- Evidence of the model :

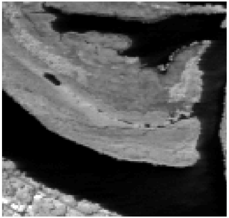
$$p(\boldsymbol{g}|\mathcal{M}) = \iint p(\boldsymbol{g}|\boldsymbol{f}, \boldsymbol{\theta}; \mathcal{M}) p(\boldsymbol{f}|\boldsymbol{\theta}; \mathcal{M}) p(\boldsymbol{\theta}|\mathcal{M})\, \mathrm{d}\boldsymbol{f}\, \mathrm{d}\boldsymbol{\theta}$$

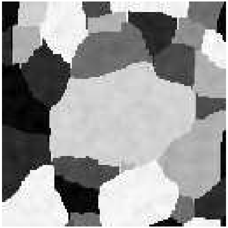
Two main steps in the Bayesian approach

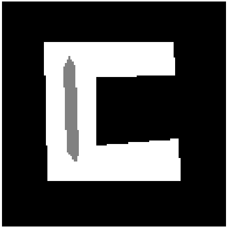
- Prior modeling
  - Separable :
    Gaussian, Generalized Gaussian, Gamma,
    mixture of Gaussians, mixture of Gammas, ...
  - Markovian : Gauss-Markov, GGM, ...
  - Separable or Markovian with hidden variables
    (contours, region labels)
- Choice of the estimator and computational aspects
  - MAP, Posterior mean, Marginal MAP
  - MAP needs optimization algorithms
  - Posterior mean needs integration methods
  - Marginal MAP needs integration and optimization
  - Approximations :
    - Gaussian approximation (Laplace)
    - Numerical exploration MCMC
    - Variational Bayes (Separable approximation)

Which images I am looking for ?

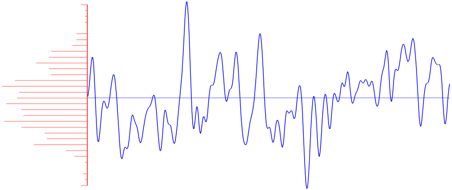




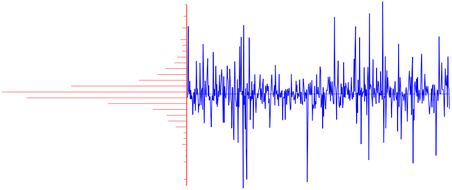




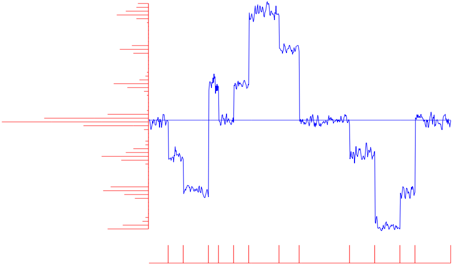
Which image I am looking for ?



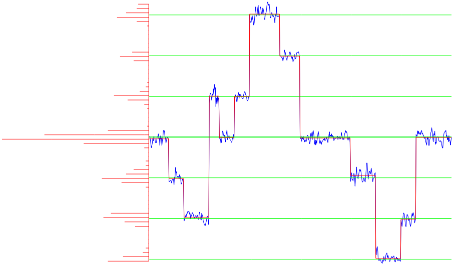
Gauss-Markov



Generalized GM



Piecewize Gaussian



Mixture of GM

Markovien prior models for images

$$\Omega(\boldsymbol{f}) = \sum_j \phi(f_j - f_{j-1})$$

- Gauss-Markov : $\phi(t) = |t|^2$
- Generalized Gauss-Markov : $\phi(t) = |t|^\alpha$
- Picewize Gauss-Markov or GGM : $\phi(t) = \begin{cases} t^2 & |t| \leq T \\ T^2 & |t| > T \end{cases}$

  or equivalently :

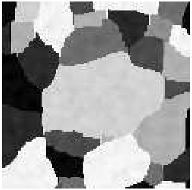
$$\Omega(\boldsymbol{f}|\boldsymbol{q}) = \sum_j (1 - q_j)\phi(f_j - f_{j-1})$$

  $\boldsymbol{q}$ line process (contours)
- Mixture of Gaussians :

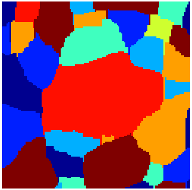
$$\Omega(\boldsymbol{f}|\boldsymbol{z}) = \sum_k \sum_{\{j:z_j=k\}} \left(\frac{f_j - m_k}{v_k}\right)^2$$

  $\boldsymbol{z}$ region labels process.

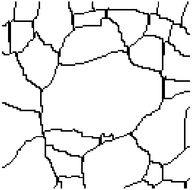
Gauss-Markov-Potts prior models for images



$f(\boldsymbol{r})$



$z(\boldsymbol{r})$



$c(\boldsymbol{r}) = 1 - \delta(z(\boldsymbol{r}) - z(\boldsymbol{r}'))$

$$p(f(\boldsymbol{r})|z(\boldsymbol{r}) = k, m_k, v_k) = \mathcal{N}(m_k, v_k)$$

$$p(f(\boldsymbol{r})) = \sum_k P(z(\boldsymbol{r}) = k)\, \mathcal{N}(m_k, v_k) \quad \text{Mixture of Gaussians}$$

- Separable iid hidden variables : $p(\boldsymbol{z}) = \prod_{\boldsymbol{r}} p(z(\boldsymbol{r}))$
- Markovian hidden variables : $p(\boldsymbol{z})$ Potts-Markov :

$$p(z(\boldsymbol{r})|z(\boldsymbol{r}'), \boldsymbol{r}' \in \mathcal{V}(\boldsymbol{r})) \propto \exp\left\{\gamma \sum_{\boldsymbol{r}' \in \mathcal{V}(\boldsymbol{r})} \delta(z(\boldsymbol{r}) - z(\boldsymbol{r}'))\right\}$$

$$p(\boldsymbol{z}) \propto \exp\left\{\gamma \sum_{\boldsymbol{r} \in \mathcal{R}} \sum_{\boldsymbol{r}' \in \mathcal{V}(\boldsymbol{r})} \delta(z(\boldsymbol{r}) - z(\boldsymbol{r}'))\right\}$$

Four different cases

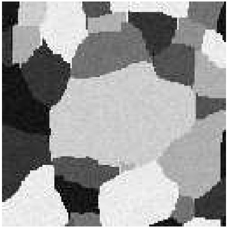
To each pixel of the image is associated 2 variables $f(\boldsymbol{r})$ and $z(\boldsymbol{r})$

- $\boldsymbol{f}|\boldsymbol{z}$ Gaussian iid, $\boldsymbol{z}$ iid :
  Mixture of Gaussians

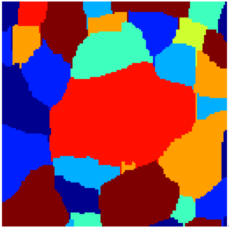
- $\boldsymbol{f}|\boldsymbol{z}$ Gauss-Markov, $\boldsymbol{z}$ iid :
  Mixture of Gauss-Markov

- $\boldsymbol{f}|\boldsymbol{z}$ Gaussian iid, $\boldsymbol{z}$ Potts-Markov :
  Mixture of Independent Gaussians
  (MIG with Hidden Potts)

- $\boldsymbol{f}|\boldsymbol{z}$ Markov, $\boldsymbol{z}$ Potts-Markov :
  Mixture of Gauss-Markov
  (MGM with hidden Potts)



$f(\boldsymbol{r})$



$z(\boldsymbol{r})$

Case 1 : $f|z$ Gaussian iid, z iid

Independent Mixture of Independent Gaussiens (IMIG) :

$$
\begin{array}{l}
p(f(\boldsymbol{r})|z(\boldsymbol{r})=k)=\mathcal{N}(m_k,v_k), \quad \forall \boldsymbol{r}\in\mathcal{R}\\
p(f(\boldsymbol{r}))=\sum_{k=1}^K \alpha_k \mathcal{N}(m_k,v_k), \text{with } \sum_k \alpha_k=1.
\end{array}
$$

$$
p(\boldsymbol{z})=\prod_{\boldsymbol{r}} p(z(\boldsymbol{r})=k)=\prod_{\boldsymbol{r}}\alpha_k=\prod_k \alpha_k^{n_k}
$$

Noting

$$
m_z(\boldsymbol{r})=m_k, v_z(\boldsymbol{r})=v_k, \alpha_z(\boldsymbol{r})=\alpha_k, \forall \boldsymbol{r}\in\mathcal{R}_k
$$

we have :

$$
\begin{array}{rcl}
p(\boldsymbol{f}|\boldsymbol{z}) & = & \prod_{\boldsymbol{r}\in\mathcal{R}} \mathcal{N}(m_z(\boldsymbol{r}),v_z(\boldsymbol{r}))\\
p(\boldsymbol{z}) & = & \prod_{\boldsymbol{r}} \alpha_z(\boldsymbol{r})=\prod_k \alpha_k^{\sum_{\boldsymbol{r}\in\mathcal{R}}\delta(z(\boldsymbol{r})-k)}=\prod_k \alpha_k^{n_k}
\end{array}
$$

Case 2 : $f|z$ Gauss-Markov, z iid

Independent Mixture of Gauss-Markov (IMGM) :

$$p(f(\boldsymbol{r})|z(\boldsymbol{r}),z(\boldsymbol{r}'),f(\boldsymbol{r}'),\boldsymbol{r}'\in\mathcal{V}(\boldsymbol{r}))=\mathcal{N}(\mu_z(\boldsymbol{r}),v_z(\boldsymbol{r})),\forall\boldsymbol{r}\in\mathcal{R}$$

$$\begin{array}{ll}
\mu_z(\boldsymbol{r}) & =\frac{1}{|\mathcal{V}(\boldsymbol{r})|}\sum_{\boldsymbol{r}'\in\mathcal{V}(\boldsymbol{r})}\mu_z^*(\boldsymbol{r}') \\
\mu_z^*(\boldsymbol{r}') & =\delta(z(\boldsymbol{r}')-z(\boldsymbol{r}))\,f(\boldsymbol{r}')+(1-\delta(z(\boldsymbol{r}')-z(\boldsymbol{r}))\,m_z(\boldsymbol{r}') \\
 & =\quad(1-c(\boldsymbol{r}'))\quad f(\boldsymbol{r}')\;+\qquad c(\boldsymbol{r}')\qquad m_z(\boldsymbol{r}')
\end{array}$$

$$\begin{array}{lll}
p(\boldsymbol{f}|\boldsymbol{z}) & \propto\prod_{\boldsymbol{r}}\mathcal{N}(\mu_z(\boldsymbol{r}),v_z(\boldsymbol{r})) & \propto\prod_k\alpha_k\mathcal{N}(m_k 1,\boldsymbol{\Sigma}_k) \\
p(\boldsymbol{z}) & =\prod_{\boldsymbol{r}}v_z(\boldsymbol{r}) & =\prod_k\alpha_k^{n_k}
\end{array}$$

with $1_k=1,\forall\boldsymbol{r}\in\mathcal{R}_k$ and $\boldsymbol{\Sigma}_k$ a covariance matrix $(n_k\times n_k)$.

Case 3 : $f|z$ Gauss iid, z Potts

Gauss iid as in Case 1 :

$$p(\boldsymbol{f}|\boldsymbol{z}) = \prod_{\boldsymbol{r}\in\mathcal{R}} \mathcal{N}(m_z(\boldsymbol{r}), v_z(\boldsymbol{r})) = \prod_k \prod_{\boldsymbol{r}\in\mathcal{R}_k} \mathcal{N}(m_k, v_k)$$

Potts-Markov

$$p(z(\boldsymbol{r})|z(\boldsymbol{r}'), \boldsymbol{r}' \in \mathcal{V}(\boldsymbol{r})) \propto \exp\left\{\gamma \sum_{\boldsymbol{r}'\in\mathcal{V}(\boldsymbol{r})} \delta(z(\boldsymbol{r}) - z(\boldsymbol{r}'))\right\}$$

$$p(\boldsymbol{z}) \propto \exp\left\{\gamma \sum_{\boldsymbol{r}\in\mathcal{R}} \sum_{\boldsymbol{r}'\in\mathcal{V}(\boldsymbol{r})} \delta(z(\boldsymbol{r}) - z(\boldsymbol{r}'))\right\}$$

Case 4 : $f|z$ Gauss-Markov, z Potts

Gauss-Markov as in Case 2 :

$$p(f(\boldsymbol{r})|z(\boldsymbol{r}), z(\boldsymbol{r}'), f(\boldsymbol{r}'), \boldsymbol{r}' \in \mathcal{V}(\boldsymbol{r})) = \mathcal{N}(\mu_z(\boldsymbol{r}), v_z(\boldsymbol{r})), \forall \boldsymbol{r} \in \mathcal{R}$$

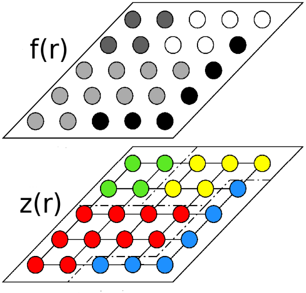
$$\begin{array}{ll} \mu_z(\boldsymbol{r}) & = \frac{1}{|\mathcal{V}(\boldsymbol{r})|} \sum_{\boldsymbol{r}' \in \mathcal{V}(\boldsymbol{r})} \mu_z^*(\boldsymbol{r}') \\ \mu_z^*(\boldsymbol{r}') & = \delta(z(\boldsymbol{r}') - z(\boldsymbol{r}))\, f(\boldsymbol{r}') + (1 - \delta(z(\boldsymbol{r}') - z(\boldsymbol{r}))\, m_z(\boldsymbol{r}') \end{array}$$

$$p(\boldsymbol{f}|\boldsymbol{z}) \quad \propto \prod_{\boldsymbol{r}} \mathcal{N}(\mu_z(\boldsymbol{r}), v_z(\boldsymbol{r})) \propto \prod_k \alpha_k \mathcal{N}(m_k 1, \boldsymbol{\Sigma}_k)$$

Potts-Markov as in Case 3 :

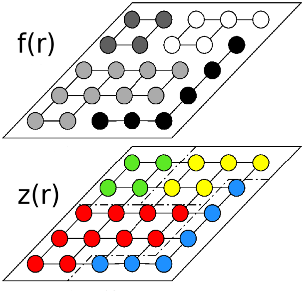
$$p(\boldsymbol{z}) \propto \exp\left\{ \gamma \sum_{\boldsymbol{r} \in \mathcal{R}} \sum_{\boldsymbol{r}' \in \mathcal{V}(\boldsymbol{r})} \delta(z(\boldsymbol{r}) - z(\boldsymbol{r}')) \right\}$$

Summary of the two proposed models



$\boldsymbol{f}|\boldsymbol{z}$ Gaussian iid
$\boldsymbol{z}$ Potts-Markov

(MIG with Hidden Potts)



$\boldsymbol{f}|\boldsymbol{z}$ Markov
$\boldsymbol{z}$ Potts-Markov

(MGM with hidden Potts)

Bayesian Computation

$$p(\boldsymbol{f},\boldsymbol{z},\boldsymbol{\theta}|\boldsymbol{g}) \propto p(\boldsymbol{g}|\boldsymbol{f},\boldsymbol{z},v_\epsilon)\, p(\boldsymbol{f}|\boldsymbol{z},\boldsymbol{m},\boldsymbol{v})\, p(\boldsymbol{z}|\gamma,\boldsymbol{\alpha})\, p(\boldsymbol{\theta})$$

$\boldsymbol{\theta} = \{v_\epsilon, (\alpha_k, m_k, v_k), k = 1, \cdot, K\}$ $\quad$ $p(\boldsymbol{\theta})$ Conjugate priors

- Direct computation and use of $p(\boldsymbol{f},\boldsymbol{z},\boldsymbol{\theta}|\boldsymbol{g};\mathcal{M})$ is too complex
- Possible approximations :
  - Gauss-Laplace (Gaussian approximation)
  - Exploration (Sampling) using MCMC methods
  - Separable approximation (Variational techniques)
- Main idea in Variational Bayesian methods : Approximate

  $p(\boldsymbol{f},\boldsymbol{z},\boldsymbol{\theta}|\boldsymbol{g};\mathcal{M})$ by $q(\boldsymbol{f},\boldsymbol{z},\boldsymbol{\theta}) = q_1(\boldsymbol{f})\, q_2(\boldsymbol{z})\, q_3(\boldsymbol{\theta})$
  - Choice of approximation criterion : $KL(q : p)$
  - Choice of appropriate families of probability laws for $q_1(\boldsymbol{f})$, $q_2(\boldsymbol{z})$ and $q_3(\boldsymbol{\theta})$

MCMC based algorithm

$$p(\boldsymbol{f},\boldsymbol{z},\boldsymbol{\theta}|\boldsymbol{g}) \propto p(\boldsymbol{g}|\boldsymbol{f},\boldsymbol{z},\boldsymbol{\theta})\,p(\boldsymbol{f}|\boldsymbol{z},\boldsymbol{\theta})\,p(\boldsymbol{z})\,p(\boldsymbol{\theta})$$

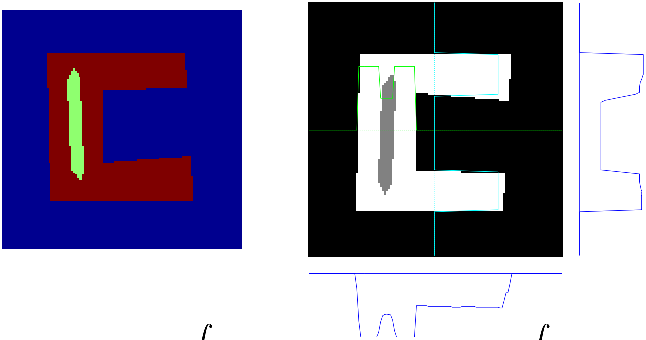
General scheme :

$$\widehat{\boldsymbol{f}} \sim p(\boldsymbol{f}|\widehat{\boldsymbol{z}},\widehat{\boldsymbol{\theta}},\boldsymbol{g}) \longrightarrow \widehat{\boldsymbol{z}} \sim p(\boldsymbol{z}|\widehat{\boldsymbol{f}},\widehat{\boldsymbol{\theta}},\boldsymbol{g}) \longrightarrow \widehat{\boldsymbol{\theta}} \sim (\boldsymbol{\theta}|\widehat{\boldsymbol{f}},\widehat{\boldsymbol{z}},\boldsymbol{g})$$

- Estimate $\boldsymbol{f}$ using $p(\boldsymbol{f}|\widehat{\boldsymbol{z}},\widehat{\boldsymbol{\theta}},\boldsymbol{g}) \propto p(\boldsymbol{g}|\boldsymbol{f},\boldsymbol{\theta})\,p(\boldsymbol{f}|\widehat{\boldsymbol{z}},\widehat{\boldsymbol{\theta}})$
  Needs optimisation of a quadratic criterion.
- Estimate $\boldsymbol{z}$ using $p(\boldsymbol{z}|\widehat{\boldsymbol{f}},\widehat{\boldsymbol{\theta}},\boldsymbol{g}) \propto p(\boldsymbol{g}|\widehat{\boldsymbol{f}},\widehat{\boldsymbol{z}},\widehat{\boldsymbol{\theta}})\,p(\boldsymbol{z})$
  Needs sampling of a Potts Markov field.
- Estimate $\boldsymbol{\theta}$ using
  $p(\boldsymbol{\theta}|\widehat{\boldsymbol{f}},\widehat{\boldsymbol{z}},\boldsymbol{g}) \propto p(\boldsymbol{g}|\widehat{\boldsymbol{f}},\sigma_\epsilon^2\boldsymbol{I})\,p(\widehat{\boldsymbol{f}}|\widehat{\boldsymbol{z}},(m_k,v_k))\,p(\boldsymbol{\theta})$
  Conjugate priors $\longrightarrow$ analytical expressions.

Application of CT in NDT

Reconstruction from only 2 projections

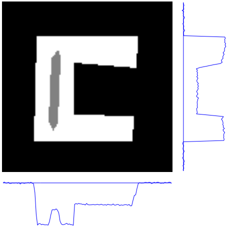
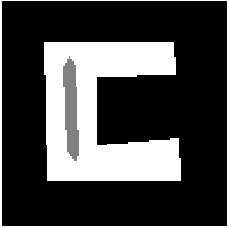
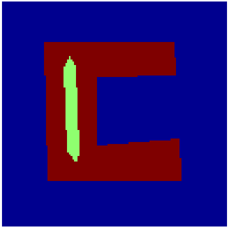
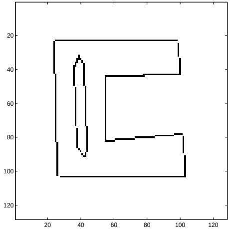


$$g_1(x) = \int f(x,y)\,\mathrm{d}y, \qquad g_2(y) = \int f(x,y)\,\mathrm{d}x$$

- Given the marginals $g_1(x)$ and $g_2(y)$ find the joint distribution $f(x,y)$.
- Infinite number of solutions : $f(x,y) = g_1(x)\,g_2(y)\,\Omega(x,y)$
  $\Omega(x,y)$ is a Copula :

$$\int \Omega(x,y)\,\mathrm{d}x = 1 \quad \text{and} \quad \int \Omega(x,y)\,\mathrm{d}y = 1$$

Application in CT

$\boldsymbol{g}|\boldsymbol{f}$

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

$\boldsymbol{g}|\boldsymbol{f} \sim \mathcal{N}(\boldsymbol{H}\boldsymbol{f}, \sigma_\epsilon^2 \boldsymbol{I})$

Gaussian

$\boldsymbol{f}|\boldsymbol{z}$

iid Gaussian

or

Gauss-Markov

$\boldsymbol{z}$

iid

or

Potts

$\boldsymbol{c}$

$c(\boldsymbol{r}) \in \{0, 1\}$

$1 - \delta(z(\boldsymbol{r}) - z(\boldsymbol{r}'))$

binary

Proposed algorithm

$$p(\boldsymbol{f},\boldsymbol{z},\boldsymbol{\theta}|\boldsymbol{g}) \propto p(\boldsymbol{g}|\boldsymbol{f},\boldsymbol{z},\boldsymbol{\theta})\, p(\boldsymbol{f}|\boldsymbol{z},\boldsymbol{\theta})\, p(\boldsymbol{\theta})$$

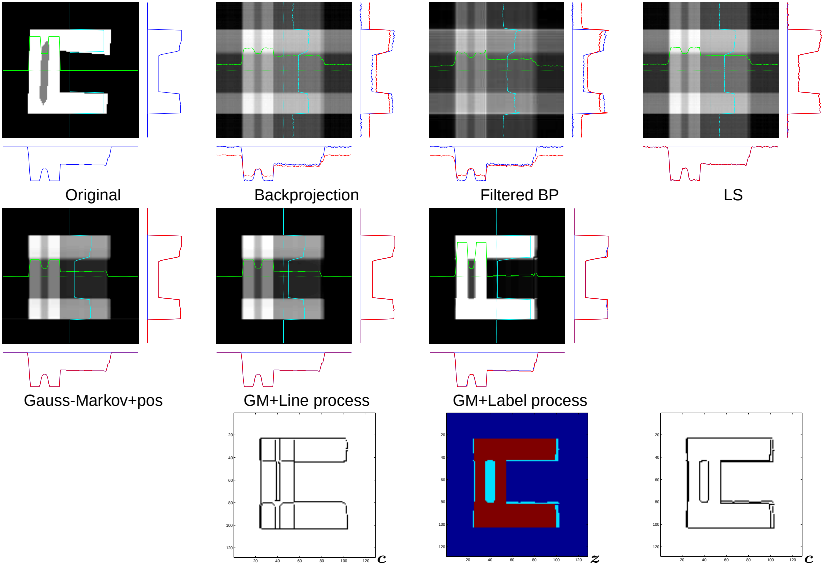
General scheme :

$$\widehat{\boldsymbol{f}} \sim p(\boldsymbol{f}|\widehat{\boldsymbol{z}},\widehat{\boldsymbol{\theta}},\boldsymbol{g}) \longrightarrow \widehat{\boldsymbol{z}} \sim p(\boldsymbol{z}|\widehat{\boldsymbol{f}},\widehat{\boldsymbol{\theta}},\boldsymbol{g}) \longrightarrow \widehat{\boldsymbol{\theta}} \sim (\boldsymbol{\theta}|\widehat{\boldsymbol{f}},\widehat{\boldsymbol{z}},\boldsymbol{g})$$

Iterative algorithme :

- Estimate $\boldsymbol{f}$ using $p(\boldsymbol{f}|\widehat{\boldsymbol{z}},\widehat{\boldsymbol{\theta}},\boldsymbol{g}) \propto p(\boldsymbol{g}|\boldsymbol{f},\boldsymbol{\theta})\, p(\boldsymbol{f}|\widehat{\boldsymbol{z}},\widehat{\boldsymbol{\theta}})$
  Needs optimisation of a quadratic criterion.
- Estimate $\boldsymbol{z}$ using $p(\boldsymbol{z}|\widehat{\boldsymbol{f}},\widehat{\boldsymbol{\theta}},\boldsymbol{g}) \propto p(\boldsymbol{g}|\widehat{\boldsymbol{f}},\widehat{\boldsymbol{z}},\widehat{\boldsymbol{\theta}})\, p(\boldsymbol{z})$
  Needs sampling of a Potts Markov field.
- Estimate $\boldsymbol{\theta}$ using
  $p(\boldsymbol{\theta}|\widehat{\boldsymbol{f}},\widehat{\boldsymbol{z}},\boldsymbol{g}) \propto p(\boldsymbol{g}|\widehat{\boldsymbol{f}},\sigma_\epsilon^2\boldsymbol{I})\, p(\widehat{\boldsymbol{f}}|\widehat{\boldsymbol{z}},(m_k,v_k))\, p(\boldsymbol{\theta})$
  Conjugate priors $\longrightarrow$ analytical expressions.

Results

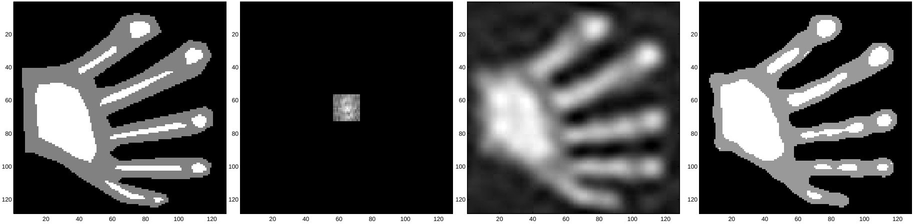

Original
Backprojection
Filtered BP
LS
Gauss-Markov+pos
GM+Line process
GM+Label process
c
z
c

Application in Microwave imaging

$$g(\boldsymbol{\omega}) = \int f(\boldsymbol{r}) \exp\{-j(\boldsymbol{\omega}.\boldsymbol{r})\}\; \mathrm{d}\boldsymbol{r} + \epsilon(\boldsymbol{\omega})$$

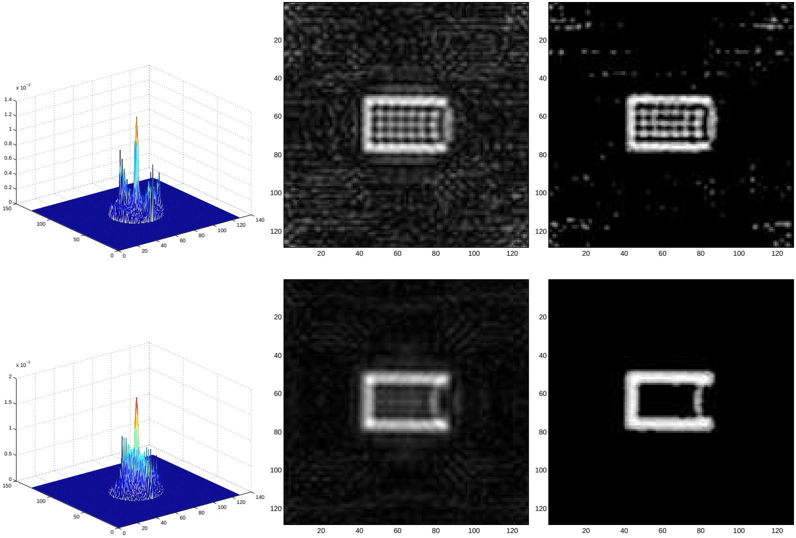
$$g(u,v) = \iint f(x,y) \exp\{-j(ux+vy)\}\; \mathrm{d}x\, \mathrm{d}y + \epsilon(u,v)$$

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



$f(x,y)$ $g(u,v)$ $\widehat{f}$ IFT $\widehat{f}$ Proposed method

Application in Microwave imaging



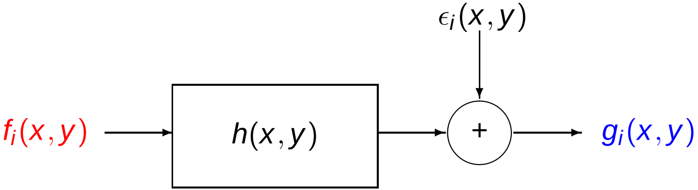
Conclusions

- Bayesian Inference for inverse problems
- Approximations (Laplace, MCMC, Variational)
- Gauss-Markov-Potts are useful prior models for images incorporating regions and contours
- Separable approximations for Joint posterior with Gauss-Markov-Potts priors
- Application in different CT (X ray, US, Microwaves, PET, SPECT)

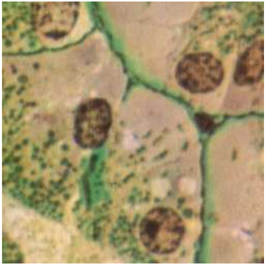
Perspectives :

- Efficient implementation in 2D and 3D cases
- Evaluation of performances and comparison with MCMC methods
- Application to other linear and non linear inverse problems :
  (PET, SPECT or ultrasound and microwave imaging)

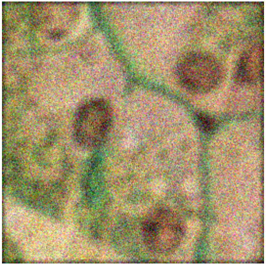
Color (Multi-spectral) image deconvolution



Observation model : $\quad \boldsymbol{g}_i = \boldsymbol{H}\boldsymbol{f}_i + \boldsymbol{\epsilon}_i, \quad i = 1, 2, 3$



$\overset{?}{\Longleftarrow}$

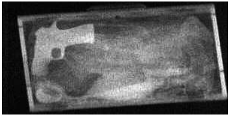


Images fusion and joint segmentation

(with O. Féron)

$$\begin{cases} g_i(\boldsymbol{r}) = f_i(\boldsymbol{r}) + \epsilon_i(\boldsymbol{r}) \\ p(f_i(\boldsymbol{r})|z(\boldsymbol{r}) = k) = \mathcal{N}(m_{i\,k}, \sigma_{i\,k}^2) \\ p(\underline{\boldsymbol{f}}|\boldsymbol{z}) = \prod_i p(\boldsymbol{f}_i|\boldsymbol{z}) \end{cases}$$

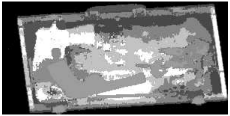
$\boldsymbol{g}_1$



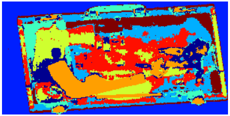
$\boldsymbol{g}_2$

$\longrightarrow$

$\widehat{\boldsymbol{f}}_1$



$\widehat{\boldsymbol{f}}_2$

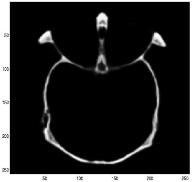


$\widehat{\boldsymbol{z}}$

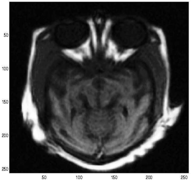
Data fusion in medical imaging

(with O. Féron)

$$\begin{cases} g_i(\boldsymbol{r}) = f_i(\boldsymbol{r}) + \epsilon_i(\boldsymbol{r}) \\ p(f_i(\boldsymbol{r})|z(\boldsymbol{r}) = k) = \mathcal{N}(m_{i\,k}, \sigma_{i\,k}^2) \\ p(\underline{\boldsymbol{f}}|\boldsymbol{z}) = \prod_i p(\boldsymbol{f}_i|\boldsymbol{z}) \end{cases}$$

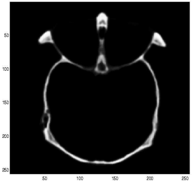


$\boldsymbol{g}_1$

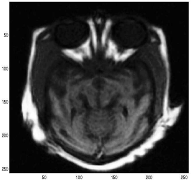


$\boldsymbol{g}_2$

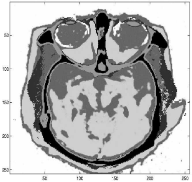
$\longrightarrow$



$\widehat{\boldsymbol{f}}_1$



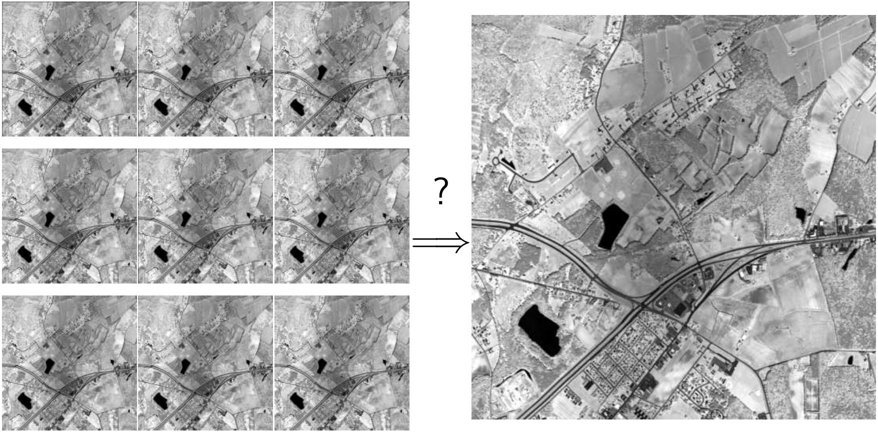
$\widehat{\boldsymbol{f}}_2$



$\widehat{\boldsymbol{z}}$

Super-Resolution

(with F. Humblot)



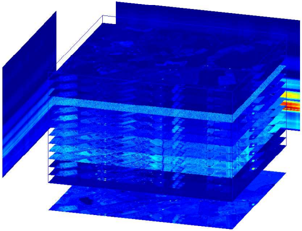
Low Resolution images

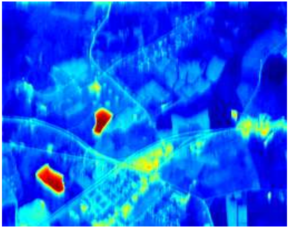
High Resolution image

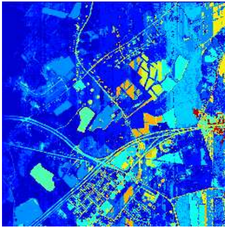
Joint segmentation of hyper-spectral images

(with N. Bali & A. Mohammadpour)

$$\begin{cases} g_i(\boldsymbol{r}) = f_i(\boldsymbol{r}) + \epsilon_i(\boldsymbol{r}) \\ p(f_i(\boldsymbol{r})|z(\boldsymbol{r})=k) = \mathcal{N}(m_{i\,k}, \sigma_{i\,k}^2), \quad k=1,\cdots,K \\ p(\underline{\boldsymbol{f}}|\boldsymbol{z}) = \prod_i p(\boldsymbol{f}_i|\boldsymbol{z}) \\ m_{i\,k} \text{ follow a Markovian model along the index } i \end{cases}$$



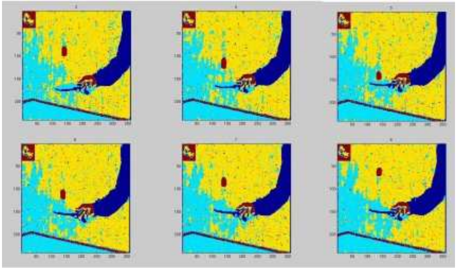




Segmentation of a video sequence of images

(with P. Brault)

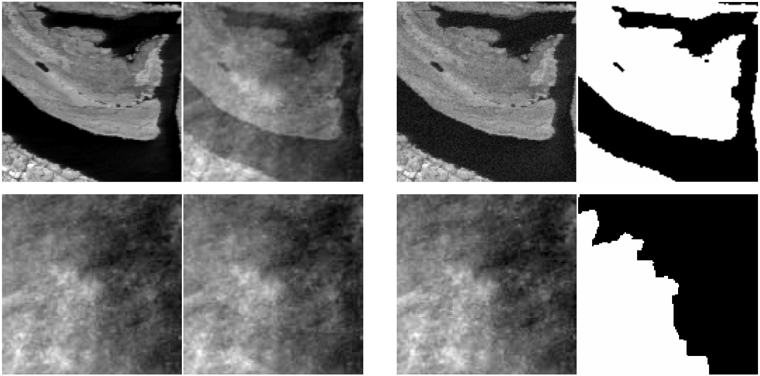
$$\left\{\begin{array}{l} g_i(\boldsymbol{r}) = f_i(\boldsymbol{r}) + \epsilon_i(\boldsymbol{r}) \\ p(f_i(\boldsymbol{r})|z_i(\boldsymbol{r}) = k) = \mathcal{N}(m_{i\,k}, \sigma_{i\,k}^2), \quad k = 1, \cdots, K \\ p(\underline{\boldsymbol{f}}|\boldsymbol{z}) = \prod_i p(\boldsymbol{f}_i|\boldsymbol{z}_i) \\ z_i(\boldsymbol{r}) \text{ follow a Markovian model along the index } i \end{array}\right.$$



Source separation

(with H. Snoussi & M. Ichir)

$$\begin{cases} g_i(\boldsymbol{r}) = \sum_{j=1}^{N} A_{ij} f_j(\boldsymbol{r}) + \epsilon_i(\boldsymbol{r}) \\ p(f_j(\boldsymbol{r}) | z_j(\boldsymbol{r}) = k) = \mathcal{N}(m_{j_k}, \sigma^2_{j\,k}) \\ p(A_{ij}) = \mathcal{N}(A_{0_{ij}}, \sigma^2_{0_{ij}}) \end{cases}$$



$\boldsymbol{f}$ $\boldsymbol{g}$ $\widehat{\boldsymbol{f}}$ $\widehat{\boldsymbol{z}}$

Some references

- O. Féron, B. Duchène and A. Mohammad-Djafari,
  Microwave imaging of inhomogeneous objects made of a finite number of dielectric and conductive materials from experimental data, Inverse Problems, 21(6) :95-115, Dec 2005.
- M. Ichir and A. Mohammad-Djafari,
  Hidden markov models for blind source separation, IEEE Trans. on Signal Processing, 15(7) :1887-1899, Jul 2006.
- F. Humblot and A. Mohammad-Djafari,
  Super-Resolution using Hidden Markov Model and Bayesian Detection Estimation Framework, EURASIP Journal on Applied Signal Processing, Special number on Super-Resolution Imaging : Analysis, Algorithms, and Applications :ID 36971, 16 pages, 2006.
- O. Féron and A. Mohammad-Djafari,
  Image fusion and joint segmentation using an MCMC algorithm, Journal of Electronic Imaging, 14(2) :paper no. 023014, Apr 2005.
- H. Snoussi and A. Mohammad-Djafari,
  Fast joint separation and segmentation of mixed images, Journal of Electronic Imaging, 13(2) :349-361, April 2004.

Questions and Discussions

- Thanks for your attentions
- ...
- ...
- Questions ?
- Discussions ?
- ...
- ...