Bayesian Computed Tomography

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

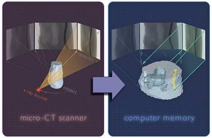
NIMI, Gent, Belgium, Mai 2010

Content

- Image reconstruction in Computed Tomography:
  An ill posed invers problem
- Two main steps in Bayesian approach:
  Prior modeling and Bayesian computation
- Prior models for images:
  - Separable Gaussian, GG, ...
  - Gauss-Markov, General one layer Markovian models
  - Hierarchical Markovian models with hidden variables (contours and regions)
  - Gauss-Markov-Potts
- Bayesian computation
  - MCMC
  - Variational and Mean Field approximations (VBA, MFA)
- Application: Computed Tomography in NDT
- Conclusions and Work in Progress
- Questions and Discussion

Computed Tomography: 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

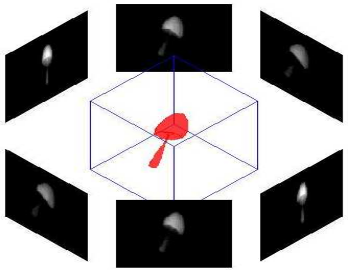
$$
\begin{aligned}
g_\phi(r) &= \int_{L_{r,\phi}} f(x, y)\,\mathrm{d}l + \epsilon_\phi(r) \\
&= \iint f(x, y)\,\delta(r - x\cos\phi - y\sin\phi)\,\mathrm{d}x\,\mathrm{d}y + \epsilon_\phi(r)
\end{aligned}
$$

- Inverse problem: Image reconstruction

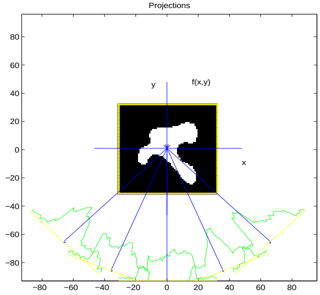
  Given the forward model $\mathcal{H}$ (Radon Transform) and
  a set of data $g_{\phi_i}(r), i = 1, \cdots, M$
  find $f(x, y)$

2D and 3D Computed Tomography

3D



2D

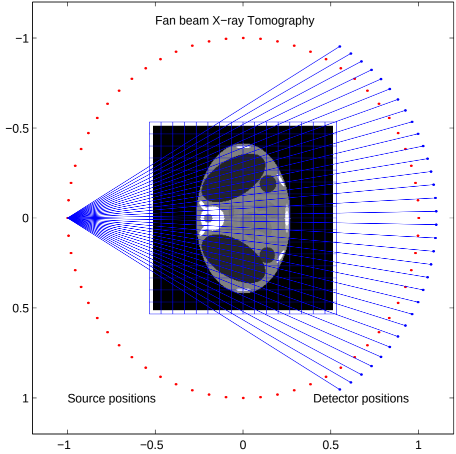


$$g_\phi(r_1, r_2) = \int_{\mathcal{L}_{r_1,r_2,\phi}} f(x,y,z) \, \mathrm{d}l \qquad g_\phi(r) = \int_{\mathcal{L}_{r,\phi}} f(x,y) \, \mathrm{d}l$$

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 $f(x,y,z)$

CT as a linear inverse problem



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

- $\boldsymbol{g}$, $\boldsymbol{f}$ and $\boldsymbol{H}$ are huge dimensional

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

– 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.

Inversion: 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

$$J(\boldsymbol{f}) = \Delta_1(\boldsymbol{g}, \boldsymbol{H}(\boldsymbol{f})) + \lambda\Delta_2(\boldsymbol{f}, \boldsymbol{f}_0)$$

Limitations:

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

Bayesian estimation approach

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

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

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 $\quad \lambda\Omega(\boldsymbol{f}) = -\ln p(\boldsymbol{f})$

But, Bayesian inference is not only limited to MAP

Case of linear models and Gaussian priors

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

- Hypothesis on the noise: $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \sigma_\epsilon^2 \boldsymbol{I})$

$$p(\boldsymbol{g}|\boldsymbol{f}) \propto \exp\left\{-\frac{1}{2\sigma_\epsilon^2}\|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{f}\|^2\right\}$$

- Hypothesis on $\boldsymbol{f}$: $\boldsymbol{f} \sim \mathcal{N}(0, \sigma_f^2 \boldsymbol{I})$

$$p(\boldsymbol{f}) \propto \exp\left\{-\frac{1}{2\sigma_f^2}\|\boldsymbol{f}\|^2\right\}$$

- A posteriori:

$$p(\boldsymbol{f}|\boldsymbol{g}) \propto \exp\left\{-\frac{1}{2\sigma_\epsilon^2}\|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{f}\|^2 - \frac{1}{2\sigma_f^2}\|\boldsymbol{f}\|^2\right\}$$

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

  with $J(\boldsymbol{f}) = \|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{f}\|^2 + \lambda\|\boldsymbol{f}\|^2, \qquad \lambda = \frac{\sigma_\epsilon^2}{\sigma_f^2}$

- Advantage : characterization of the solution

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

MAP estimation with other priors:

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

Separable priors:

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

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

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

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

- 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: $\qquad (\widehat{\boldsymbol{f}}, \widehat{\boldsymbol{\theta}}) = \arg\max_{(\boldsymbol{f},\boldsymbol{\theta})} \{p(\boldsymbol{f}, \boldsymbol{\theta}|\boldsymbol{g}; \mathcal{M})\}$
- Marginalization: $\left\{ \begin{array}{ll} p(\boldsymbol{f}|\boldsymbol{g}; \mathcal{M}) & = \int p(\boldsymbol{f}, \boldsymbol{\theta}|\boldsymbol{g}; \mathcal{M})\, \mathrm{d}\boldsymbol{f} \\ p(\boldsymbol{\theta}|\boldsymbol{g}; \mathcal{M}) & = \int p(\boldsymbol{f}, \boldsymbol{\theta}|\boldsymbol{g}; \mathcal{M})\, \mathrm{d}\boldsymbol{\theta} \end{array} \right.$
- Posterior means: $\left\{ \begin{array}{ll} \widehat{\boldsymbol{f}} & = \int \boldsymbol{f}\; p(\boldsymbol{f}, \boldsymbol{\theta}|\boldsymbol{g}; \mathcal{M})\, \mathrm{d}\boldsymbol{f}\; \mathrm{d}\boldsymbol{\theta} \\ \widehat{\boldsymbol{\theta}} & = \int \boldsymbol{\theta}\; p(\boldsymbol{f}, \boldsymbol{\theta}|\boldsymbol{g}; \mathcal{M})\, \mathrm{d}\boldsymbol{f}\; \mathrm{d}\boldsymbol{\theta} \end{array} \right.$
- Evidence of the model:

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

MAP estimation with different prior models

$$\begin{cases} \boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\epsilon} \\ p(\boldsymbol{f}) = \mathcal{N}\left(0, \sigma_f^2 (\boldsymbol{D}^t\boldsymbol{D})^{-1}\right) \end{cases} = \begin{cases} \boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\epsilon} \\ \boldsymbol{f} = \boldsymbol{C}\boldsymbol{f} + \boldsymbol{z} \quad \text{with} \quad \boldsymbol{z} \sim \mathcal{N}(0, \sigma_f^2 \boldsymbol{I}) \\ \boldsymbol{D}\boldsymbol{f} = \boldsymbol{z} \text{ with } \boldsymbol{D} = (\boldsymbol{I} - \boldsymbol{C}) \end{cases}$$

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

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

$$\begin{cases} \boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\epsilon} \\ p(\boldsymbol{f}) = \mathcal{N}\left(0, \sigma_f^2 (\boldsymbol{W}\boldsymbol{W}^t)\right) \end{cases} = \begin{cases} \boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\epsilon} \\ \boldsymbol{f} = \boldsymbol{W}\boldsymbol{z} \quad \text{with} \quad \boldsymbol{z} \sim \mathcal{N}(0, \sigma_f^2 \boldsymbol{I}) \end{cases}$$

$$p(\boldsymbol{z}|\boldsymbol{g}) = \mathcal{N}(\widehat{\boldsymbol{z}}, \widehat{\boldsymbol{P}}_z) \text{ with } \widehat{\boldsymbol{z}} = \widehat{\boldsymbol{P}}_z \boldsymbol{W}^t\boldsymbol{H}^t\boldsymbol{g}, \ \widehat{\boldsymbol{P}}_z = \left(\boldsymbol{W}^t\boldsymbol{H}^t\boldsymbol{H}\boldsymbol{W} + \lambda \boldsymbol{I}\right)^{-1}$$

$$J(\boldsymbol{z}) = -\ln p(\boldsymbol{z}|\boldsymbol{g}) = \|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{W}\boldsymbol{z}\|^2 + \lambda\|\boldsymbol{z}\|^2 \longrightarrow \widehat{\boldsymbol{f}} = \boldsymbol{W}\widehat{\boldsymbol{z}}$$

$\boldsymbol{z}$ decomposition coefficients

MAP estimation and Compressed Sensing

$$\left\{ \begin{array}{l} \boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\epsilon} \\ \boldsymbol{f} = \boldsymbol{W}\boldsymbol{z} \end{array} \right.$$

- $\boldsymbol{W}$ a code book matrix, $\boldsymbol{z}$ coefficients
- Gaussian:

$$\begin{array}{l} p(\boldsymbol{z}) = \mathcal{N}(0, \sigma_z^2 \boldsymbol{I}) \propto \exp\left\{ -\frac{1}{2\sigma_z^2} \sum_j |z_j|^2 \right\} \\ J(\boldsymbol{z}) = -\ln p(\boldsymbol{z}|\boldsymbol{g}) = \|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{W}\boldsymbol{z}\|^2 + \lambda \sum_j |z_j|^2 \end{array}$$

- Generalized Gaussian (sparsity, $\beta = 1$):

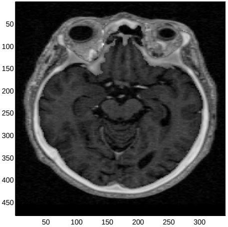
$$\begin{array}{l} p(\boldsymbol{z}) \propto \exp\left\{ -\lambda \sum_j |z_j|^\beta \right\} \\ J(\boldsymbol{z}) = -\ln p(\boldsymbol{z}|\boldsymbol{g}) = \|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{W}\boldsymbol{z}\|^2 + \lambda \sum_j |z_j|^\beta \end{array}$$

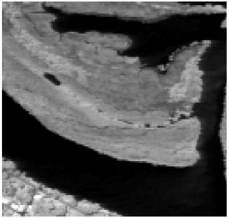
- $\boldsymbol{z} = \arg\min_{\boldsymbol{z}} \{J(\boldsymbol{z})\} \longrightarrow \widehat{\boldsymbol{f}} = \boldsymbol{W}\widehat{\boldsymbol{z}}$

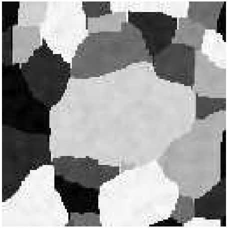
Two main steps in the Bayesian approach

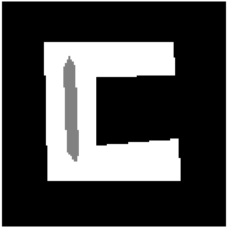
- Prior modeling
  - Separable:
    Gaussian, Generalized Gaussian, Gamma,
    mixture of Gaussians, mixture of Gammas, ...
  - Markovian: Gauss-Markov, GGM, ...
  - Separable or Markovian with hidden variables
    (contours, region labels)
- Choice of the estimator and computational aspects
  - MAP, 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?

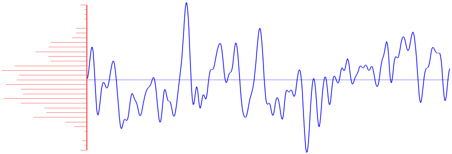






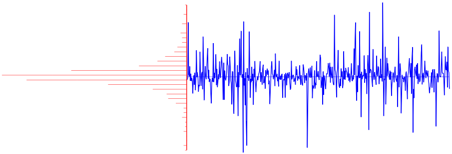


Which image I am looking for?



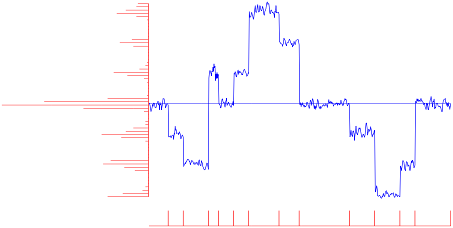
Gaussian

$p(f_j|f_{j-1}) \propto \exp\left\{-\alpha|f_j - f_{j-1}|^2\right\}$



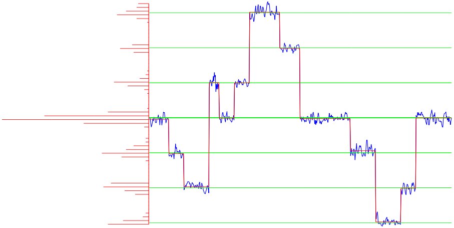
Generalized Gaussian

$p(f_j|f_{j-1}) \propto \exp\left\{-\alpha|f_j - f_{j-1}|^p\right\}$



Piecewize Gaussian

$p(f_j|q_j, f_{j-1}) = \mathcal{N}\left((1-q_j)f_{j-1}, \sigma_f^2\right)$



Mixture of GM

$p(f_j|z_j = k) = \mathcal{N}\left(m_k, \sigma_k^2\right)$

Gauss-Markov-Potts prior models for images

"In NDT applications of CT, the objects are, in general, composed of a finite number of materials, and the voxels corresponding to each materials are grouped in compact regions"

How to model this prior information?

$f(\boldsymbol{r})$

$z(\boldsymbol{r}) \in \{1, ..., K\}$

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

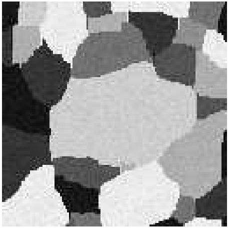
$$p(f(\boldsymbol{r})) = \sum_k P(z(\boldsymbol{r}) = k)\,\mathcal{N}(m_k, v_k) \quad \text{Mixture of Gaussians}$$

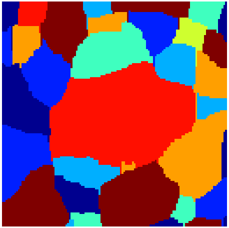
$$p(z(\boldsymbol{r})|z(\boldsymbol{r}'), \boldsymbol{r}' \in \mathcal{V}(\boldsymbol{r})) \propto \exp\left\{\gamma \sum_{\boldsymbol{r}' \in \mathcal{V}(\boldsymbol{r})} \delta(z(\boldsymbol{r}) - z(\boldsymbol{r}'))\right\}$$

Four different cases

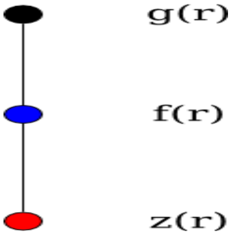
To each pixel of the image is associated 2 variables $f(r)$ and $z(r)$

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

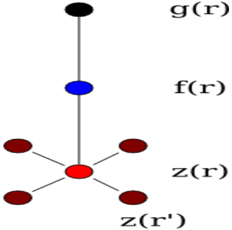

$f(\boldsymbol{r})$


$z(\boldsymbol{r})$

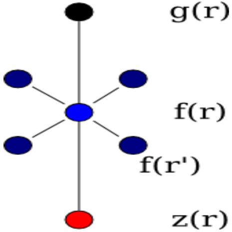
Four different cases



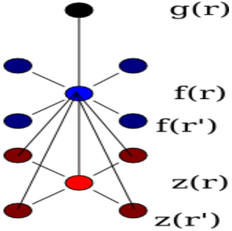
Case 1: Mixture of Gaussians



Case 2: Mixture of Gauss-Markov

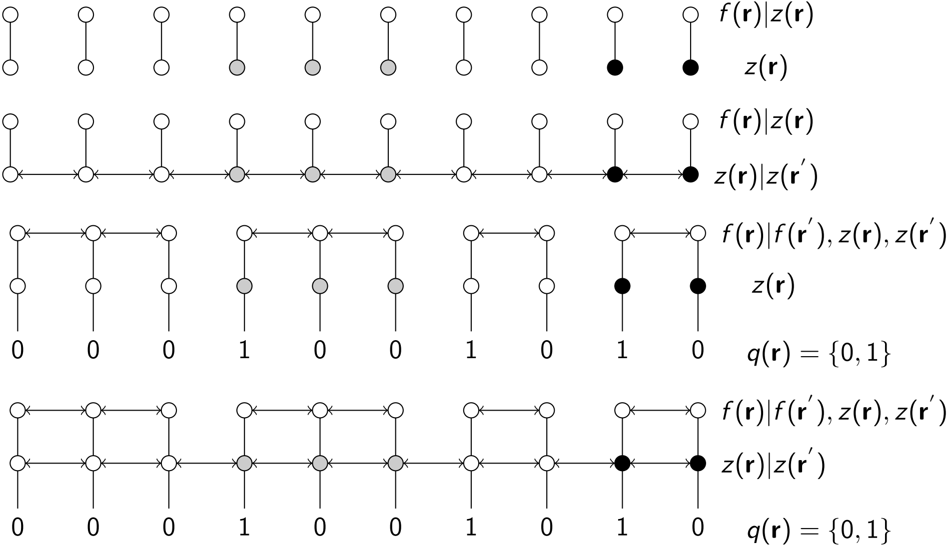


Case 3: MIG with Hidden Potts



Case 4: MGM with hidden Potts

Four different cases



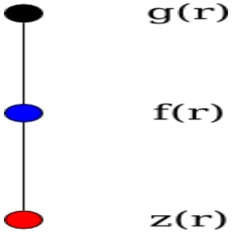
Case 1: $f|z$ Gaussian iid, z iid

Independent Mixture of Independent Gaussiens (IMIG):

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

$$p(f(\boldsymbol{r}))=\sum_{k=1}^K \alpha_k \mathcal{N}(m_k,v_k), \text{with } \sum_k \alpha_k=1.$$

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



Noting $\quad \mathcal{R}_k=\{\boldsymbol{r}: z(\boldsymbol{r})=k\}, \quad \mathcal{R}=\cup_k \mathcal{R}_k,$

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

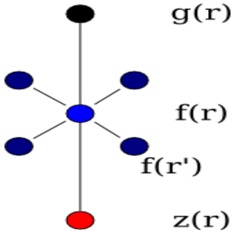
we have:

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

$$p(\boldsymbol{z}) = \prod_{\boldsymbol{r}} \alpha_z(\boldsymbol{r}) = \prod_k \alpha_k^{\sum_{\boldsymbol{r}\in\mathcal{R}} \delta(z(\boldsymbol{r})-k)} = \prod_k \alpha_k^{n_k}$$

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

Independent Mixture
of Gauss-Markov (IMGM):



$$p(f(\boldsymbol{r})|z(\boldsymbol{r}), z(\boldsymbol{r}'), f(\boldsymbol{r}'), \boldsymbol{r}' \in \mathcal{V}(\boldsymbol{r}))$$

$$= \mathcal{N}(\mu_z(\boldsymbol{r}), v_z(\boldsymbol{r})), \forall \boldsymbol{r} \in \mathcal{R}$$

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

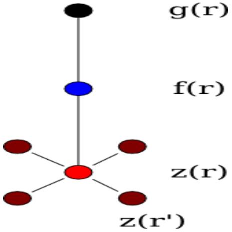
$$\begin{array}{lll}
p(\boldsymbol{f}|\boldsymbol{z}) & \propto \prod_{\boldsymbol{r}} \mathcal{N}(\mu_z(\boldsymbol{r}), v_z(\boldsymbol{r})) & \propto \prod_k \alpha_k \mathcal{N}(m_k 1, \boldsymbol{\Sigma}_k) \\
p(\boldsymbol{z}) & = \prod_{\boldsymbol{r}} v_z(\boldsymbol{r}) & = \prod_k \alpha_k^{n_k}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

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:

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



Potts-Markov:

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

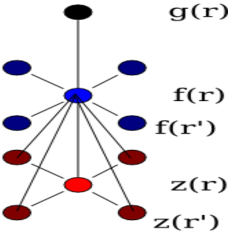
$$p(\boldsymbol{z}) \propto \exp\left\{\gamma \sum_{\boldsymbol{r}\in\mathcal{R}} \sum_{\boldsymbol{r}'\in\mathcal{V}(\boldsymbol{r})} \delta(z(\boldsymbol{r}) - z(\boldsymbol{r}'))\right\}$$

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

Gauss-Markov as in Case 2:

$$p(f(\boldsymbol{r})|z(\boldsymbol{r}), z(\boldsymbol{r}'), f(\boldsymbol{r}'), \boldsymbol{r}' \in \mathcal{V}(\boldsymbol{r})) =$$

$$\mathcal{N}(\mu_z(\boldsymbol{r}), v_z(\boldsymbol{r})), \forall \boldsymbol{r} \in \mathcal{R}$$



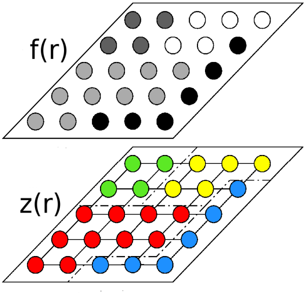
$$\begin{aligned} \mu_z(\boldsymbol{r}) &= \frac{1}{|\mathcal{V}(\boldsymbol{r})|} \sum_{\boldsymbol{r}' \in \mathcal{V}(\boldsymbol{r})} \mu_z^*(\boldsymbol{r}') \\ \mu_z^*(\boldsymbol{r}') &= \delta(z(\boldsymbol{r}') - z(\boldsymbol{r}))\, f(\boldsymbol{r}') + (1 - \delta(z(\boldsymbol{r}') - z(\boldsymbol{r}))\,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

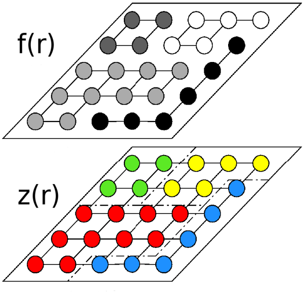
$$p(\boldsymbol{f}|\boldsymbol{z}) \quad \propto \prod_{\boldsymbol{r}} \mathcal{N}(\mu_z(\boldsymbol{r}), v_z(\boldsymbol{r})) \propto \prod_k \alpha_k \mathcal{N}(m_k 1, \boldsymbol{\Sigma}_k)$$

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

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

(MIG with Hidden Potts) (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

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

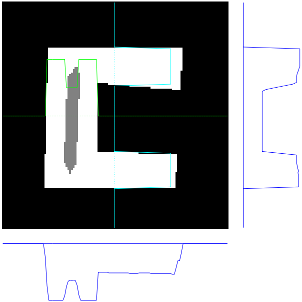
General scheme:

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

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

Application of CT in NDT

Reconstruction from only 2 projections



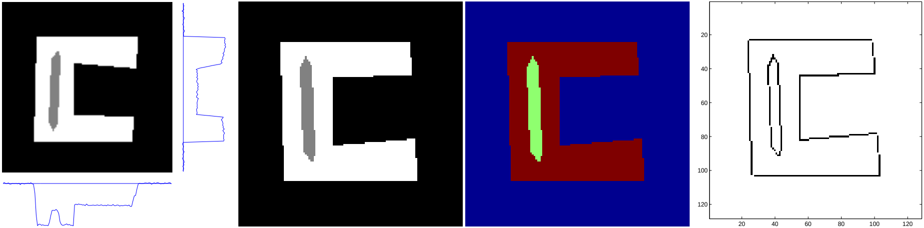
$$g_1(x) = \int f(x,y)\, \mathrm{d}y$$

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

- 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



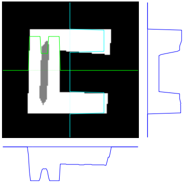
$\boldsymbol{g}\|\boldsymbol{f}$	$\boldsymbol{f}\|\boldsymbol{z}$	$\boldsymbol{z}$	$\boldsymbol{q}$
$\boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\epsilon}$	iid Gaussian	iid	$q(\boldsymbol{r}) \in \{0, 1\}$
$\boldsymbol{g}\|\boldsymbol{f} \sim \mathcal{N}(\boldsymbol{H}\boldsymbol{f}, \sigma_\epsilon^2 \boldsymbol{I})$	or	or	$1 - \delta(z(\boldsymbol{r}) - z(\boldsymbol{r}'))$
Gaussian	Gauss-Markov	Potts	binary

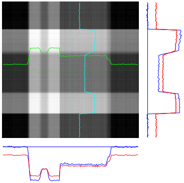
Forward model | Gauss-Markov-Potts Prior Model | Auxilary

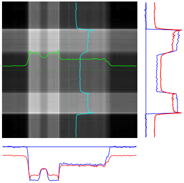
Unsupervised Bayesian estimation:

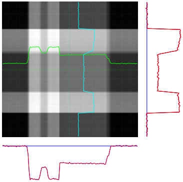
$$p(\boldsymbol{f}, \boldsymbol{z}, \boldsymbol{\theta}|\boldsymbol{g}) \propto p(\boldsymbol{g}|\boldsymbol{f}, \boldsymbol{z}, \boldsymbol{\theta})\, p(\boldsymbol{f}|\boldsymbol{z}, \boldsymbol{\theta})\, p(\boldsymbol{\theta})$$

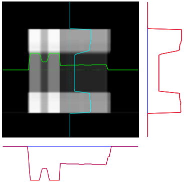
Results: 2D case

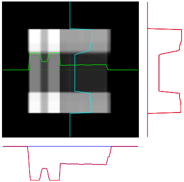

Original

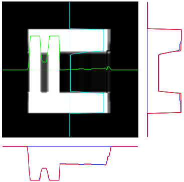

Backprojection

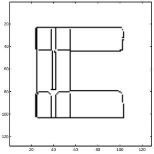

Filtered BP

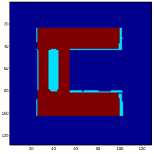

LS

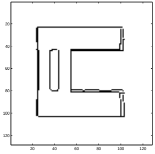

Gauss-Markov+pos


GM+Line process


GM+Label process

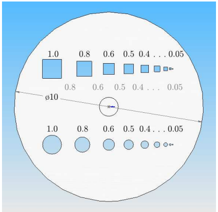

c


z

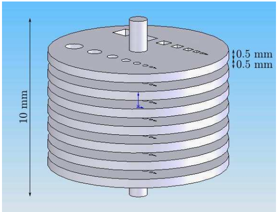

c

Some results in 3D case

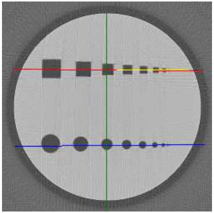
(Results obtained with collaboration with CEA)



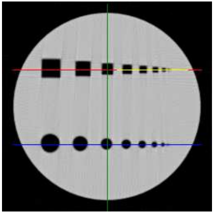
M. Defrise



Phantom

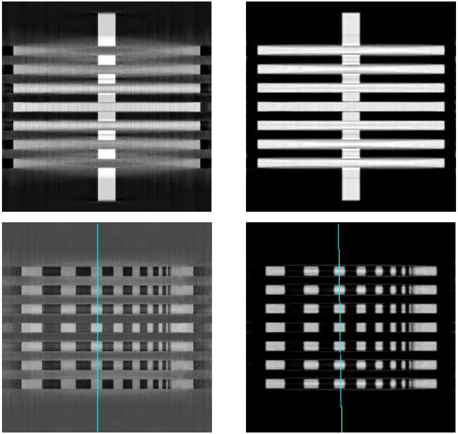


FeldKamp



Proposed method

Some results in 3D case

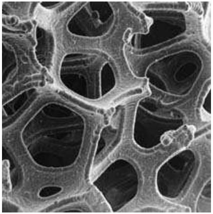


FeldKamp   Proposed method

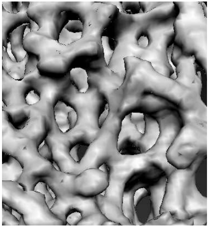
Some results in 3D case

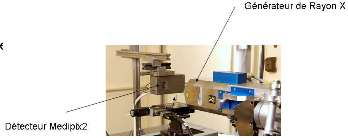
Experimental setup

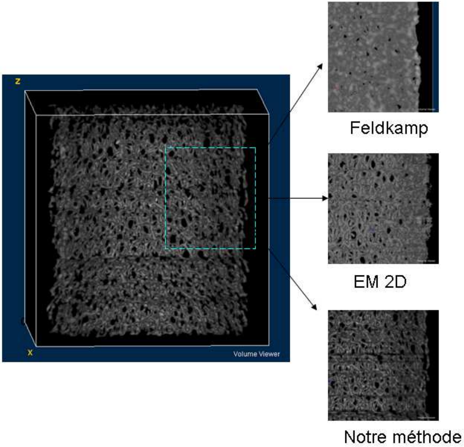
A photograpy of metalique esponge



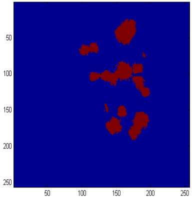
Reconstruction by proposed method



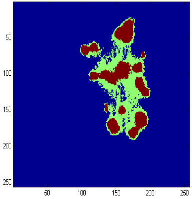




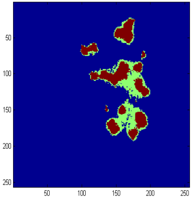
Application: liquid evaporation in metalic esponge



Time 0



Time 1



Time 2

Conclusions

- Gauss-Markov-Potts are useful prior models for images incorporating regions and contours
- Bayesian computation needs often pproximations (Laplace, MCMC, Variational Bayes)
- Application in different CT systems (X ray, Ultrasound, Microwave, PET, SPECT) as well as other inverse problems

Work in Progress and Perspectives :

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

Some references

- A. Mohammad-Djafari (Ed.) Problèmes inverses en imagerie et en vision (Vol. 1 et 2), Hermes-Lavoisier, Traité Signal et Image, IC2, 2009,
- A. Mohammad-Djafari (Ed.) Inverse Problems in Vision and 3D Tomography, ISTE, Wiley and sons, ISBN: 9781848211728, December 2009, Hardback, 480 pp.
- H. Ayasso and Ali Mohammad-Djafari Joint NDT Image Restoration and Segmentation using Gauss-Markov-Potts Prior Models and Variational Bayesian Computation, To appear in IEEE Trans. on Image Processing, TIP-04815-2009.R2, 2010.
- H. Ayasso, B. Duchene and A. Mohammad-Djafari, Bayesian Inversion for Optical Diffraction Tomography Journal of Modern Optics, 2008.
- A. Mohammad-Djafari, Gauss-Markov-Potts Priors for Images in Computer Tomography Resulting to Joint Optimal Reconstruction and segmentation, International Journal of Tomography & Statistics 11: W09. 76-92, 2008.
- A Mohammad-Djafari, Super-Resolution : A short review, a new method based on hidden Markov modeling of HR image and future challenges, The Computer Journal doi:10,1093/comjnl/bxn005:, 2008.
- 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.
- A. Mohammad-Djafari, J.F. Giovannelli, G. Demoment and J. Idier, Regularization, maximum entropy and probabilistic methods in mass spectrometry data processing problems, Int. Journal of Mass Spectrometry, 215(1-3):175-193, April 2002.

Thanks, Questions and Discussions

Thanks to:

My graduated PhD students:

- H. Snoussi, M. Ichir, (Sources separation)
- F. Humblot (Super-resolution)
- H. Carfantan, O. Féron (Microwave Tomography)
- S. Fékih-Salem (3D X ray Tomography)

My present PhD students:

- H. Ayasso (Optical Tomography, Variational Bayes)
- D. Pougaza (Tomography and Copula)
- ___________
- Sh. Zhu (SAR Imaging)
- D. Fall (Emission Positon Tomography, Non Parametric Bayesian)

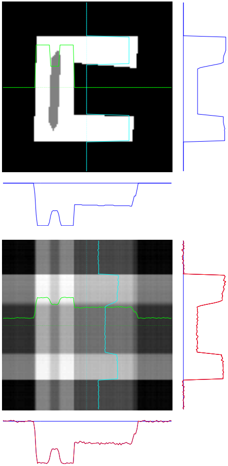
My colleages in GPI (L2S) & collaborators in other instituts:

- B. Duchêne & A. Joisel (Inverse scattering and Microwave Imaging)
- N. Gac & A. Rabanal (GPU Implementation)
- Th. Rodet (Tomography)
- ___________
- A. Vabre & S. Legoupil (CEA-LIST), (3D X ray Tomography)
- E. Barat (CEA-LIST) (Positon Emission Tomography, Non Parametric Bayesian)
- C. Comtat (SHFJ, CEA)(PET, Spatio-Temporal Brain activity)

Questions and Discussions

Continuous Signals and Images
Simple Gauss-Markov model

$g = H f + \epsilon$
$f(r)$ Continuous Image : Gauss-Markov
$p(f) = N(0, \Sigma_f)$
$p(f_j | f_i, i \neq j) = \mathcal{N}(\beta f_{j-1}, \sigma_f^2)$
$p(f(r) | f(s)) = \mathcal{N}\left(\beta \sum_{s \in \mathcal{V}(r)} f(s), \sigma_f^2\right)$



MAP :
$\widehat{f} = \arg\max_f \{p(f|g)\} = \arg\min_f \{J(f)\}$

$J(f) = \|g - H f\|^2 + \sum_j (f_j - \beta f_{j-1})^2$

$$J(f) = \|g - H f\|^2 + \sum_r \left(f(r) - \beta \sum_{s \in \mathcal{V}(r)} f(s)\right)^2$$

Piecewise continuous signals and images

Compound Gauss-Markov with contours hidden variable model

$f(\boldsymbol{r})$ piecewise continuous image :

Compound Intensity-Contours MRF

Hidden contour variable: $q(\boldsymbol{r})$

$p(f_j|q_j, f_i, i \neq j) = \mathcal{N}(\beta(1-q_j) f_{j-1}, \sigma_f^2)$

$p(f(\boldsymbol{r})|q(\boldsymbol{r}), f(\boldsymbol{s}))$

$= \mathcal{N}\left(\beta(1-q(\boldsymbol{r}))\sum_{\boldsymbol{s}\in\mathcal{V}(\boldsymbol{r})} f(\boldsymbol{s}), \sigma_f^2\right)$

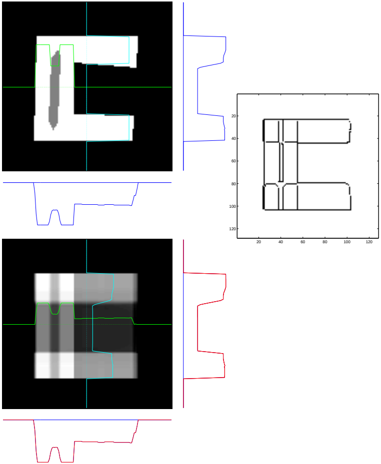
MAP : $(\widehat{\boldsymbol{f}}, \widehat{\boldsymbol{q}}) = \arg\max_{\boldsymbol{f},\boldsymbol{q}} \{p(\boldsymbol{f}, \boldsymbol{q}|\boldsymbol{g})\}$

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

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

$$+ \sum_{\boldsymbol{r}} (1 - q(\boldsymbol{r})) \left(f(\boldsymbol{r}) - \beta \sum_{\boldsymbol{s}\in\mathcal{V}(\boldsymbol{r})} f(\boldsymbol{s}) \right)^2$$

$$\widehat{\boldsymbol{q}} = \arg\max_{\boldsymbol{q}} \{p(\boldsymbol{q}|\boldsymbol{g})\}$$



Objects composed of a finite number of homogeneous materials

$\boldsymbol{f}$: Compound intensity-regions MRF

Introduction of a class label variable $z(\boldsymbol{r})$

$z(\boldsymbol{r}) = k, \quad k = 1, \cdots, K$

$\mathcal{R}_k = \{\boldsymbol{r} \ : \ z(\boldsymbol{r}) = k\}, \quad \mathcal{R} = \cup_k \mathcal{R}_k$

$p(f(\boldsymbol{r})|z(\boldsymbol{r}) = k) = \mathcal{N}(f(\boldsymbol{r})|m_k, \sigma_k^2)$

$\boldsymbol{z} = \{z(\boldsymbol{r}), \boldsymbol{r} \in \mathcal{R}\}$ a segmented image

Potts MRF:

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

$$p(\boldsymbol{f}, \boldsymbol{z}, \boldsymbol{\theta}|\boldsymbol{g}) \propto p(\boldsymbol{g}|\boldsymbol{f}, \boldsymbol{\theta}_1)\ p(\boldsymbol{f}|\boldsymbol{z}, \boldsymbol{\theta}_2)\ p(\boldsymbol{z}|\boldsymbol{\theta}_2)\ p(\boldsymbol{\theta})$$

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

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

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