Gauss-Markov-Potts Priors for Inverse Problems in Imaging Systems

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

Séminaires du Centre Mathématique de Jussieu-Chevaleret,
25/03/2009
Computed Tomography (CT) as an Invers Problem example
Classical methods: analytical and algebraic method
Probabilistic methods
Bayesian inference approach
Gauss-Markov-Potts prior models for images
Bayesian computation
VB with Gauss-Markov-Potts prior models
Application in Computed Tomography
Conclusions
Questions and Discussion
2D and 3D Computed Tomography

\[ g_\phi(r_1, r_2) = \int_{\mathcal{L}_{r_1, r_2, \phi}} f(x, y, z) \, dl \quad g_\phi(r) = \int_{\mathcal{L}_{r, \phi}} f(x, y) \, dl \]

Forward problem: \( f(x, y) \) or \( f(x, y, z) \) \( \rightarrow \) \( g_\phi(r) \) or \( g_\phi(r_1, r_2) \)

Inverse problem: \( g_\phi(r) \) or \( g_\phi(r_1, r_2) \) \( \rightarrow \) \( f(x, y) \) or \( f(x, y, z) \)
X ray Tomography and Radon Transform

\[ g(r, \phi) = -\ln \left( \frac{l}{l_0} \right) = \int_{L_{r,\phi}} f(x, y) \, dl \]

\[ g(r, \phi) = \iiint_D f(x, y) \delta(r - x \cos \phi - y \sin \phi) \, dx \, dy \]

\[ f(x, y) \xrightarrow{RT} g(r, \phi) \]

IRT

\[ \Rightarrow \]
Analytical Inversion methods

Radon:

\[ g(r, \phi) = \int_L f(x, y) \, dl \]

\[ f(x, y) = \left( -\frac{1}{2\pi^2} \right) \int_0^\pi \int_{-\infty}^{+\infty} \frac{\partial}{\partial r} g(r, \phi) \frac{dr \, d\phi}{(r - x \cos \phi - y \sin \phi)} \]
Filtered Backprojection method

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

Derivation \( \mathcal{D} \) :
\[ \bar{g}(r, \phi) = \frac{\partial g(r, \phi)}{\partial r} \]

Hilbert Transform \( \mathcal{H} \) :
\[ g_1(r', \phi) = \frac{1}{\pi} \int_0^\infty \frac{\bar{g}(r, \phi)}{r - r'} \, dr \]

Backprojection \( \mathcal{B} \) :
\[ f(x, y) = \frac{1}{2\pi} \int_0^\pi g_1(r' = x \cos \phi + y \sin \phi, \phi) \, d\phi \]

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

- Backprojection of filtered projections:

\[ \begin{array}{c|c|c|c|c}
 g(r, \phi) & \text{FT} & \text{Filter} & \text{IFT} & f(x, y) \\
 \mathcal{F}_1 & |\Omega| & \mathcal{F}_1^{-1} & \mathcal{B} & \\
 \end{array} \]
Limitations: Limited angle or noisy data

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

Original Data | Backprojection | Filtered Backprojection
---|---|---
CT as a linear inverse problem

\[ g(s_i) = \int_{L_i} f(r) \, dl_i \quad \rightarrow \quad \text{Discretization} \quad \rightarrow \quad g = Hf + \epsilon \]
Classical methods in CT

\[ g(s_i) = \int_{L_i} f(r) \, dl_i \quad \text{Discretization} \quad \rightarrow \quad g = Hf + \epsilon \]

- **H** is a huge dimensional matrix of line integrals
- **Hf** is the forward or projection operation
- **H^t g** is the backward or backprojection operation
- \((H^t H)^{-1} H^t g\) is the filtered backprojection minimizing \(\|g - Hf\|^2\)
- Iterative methods:
  \[ \hat{f}^{(k+1)} = \hat{f}^{(k)} + \alpha^{(k)} H^t (g - H \hat{f}^{(k)}) \]
  is the Least squares iterative reconstruction method
- Regularization:
  \[ J(f) = \|g - Hf\|^2 + \lambda \|Df\|^2. \]
Inversion : Deterministic methods

Data matching

- Observation model
  \[ g_i = h_i(f) + \epsilon_i, \quad i = 1, \ldots, M \rightarrow g = H(f) + \epsilon \]
- Misatch between data and output of the model \( \Delta(g, H(f)) \)
  \[ \hat{f} = \arg \min_{f} \{ \Delta(g, H(f)) \} \]
- Examples:
  - LS \( \Delta(g, H(f)) = \| g - H(f) \|_2^2 = \sum_i |g_i - h_i(f)|^2 \)
  - \( L_p \) \( \Delta(g, H(f)) = \| g - H(f) \|_p^p = \sum_i |g_i - h_i(f)|^p, \quad 1 < p < 2 \)
  - KL \( \Delta(g, H(f)) = \sum_i g_i \ln \frac{g_i}{h_i(f)} \)

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

Inverse problems = Ill posed problems
  \[ \implies \text{Need of prior information} \]

- Minimum norme LS (MNLS) :
  \[ J(f) = \|g - H(f)\|^2 + \lambda \|f\|^2 \]

- Classical regularization :
  \[ J(f) = \|g - H(f)\|^2 + \lambda \|Df\|^2 \]
  \[ = \|g - H(f)\|^2 + \lambda \sum_j \|Df_j\|^2 \]

- More general regularization :
  \[ J(f) = Q(g - H(f)) + \lambda \sum_j \Phi([Df]_j) \]
  or \[ J(f) = \Delta_1(g, H(f)) + \lambda \Delta_2(f, f_\infty) \]

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 $\rightarrow$ 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

\[ g = Hf + \epsilon \]

- Observation model \( \mathcal{M} \) + Hypothesis on the noise \( \epsilon \)
  \[ \rightarrow p(g | f; \mathcal{M}) = p_\epsilon (g - Hf) \]

- A priori information \( p(f | \mathcal{M}) \)

- Bayes:
  \[ p(f | g; \mathcal{M}) = \frac{p(g | f; \mathcal{M}) p(f | \mathcal{M})}{p(g | \mathcal{M})} \]

Link with regularisation:
Maximum A Posteriori (MAP):

\[ \hat{f} = \arg \max_f \{ p(f | g) \} = \arg \max_f \{ p(g | f) p(f) \} \]

\[ \hat{f} = \arg \min_f \{ - \ln p(g | f) - \ln p(f) \} \]

with

\[ Q(g, Hf) = - \ln p(g | f) \quad \text{and} \quad \lambda \Phi(f) = - \ln p(f) \]
Case of linear models and Gaussian priors

\[ g = Hf + \epsilon \]

- Hypothesis on the noise: \( \epsilon \sim \mathcal{N}(0, \sigma^2_\epsilon I) \)
  \[ p(g|f) \propto \exp \left\{ -\frac{1}{2\sigma^2_\epsilon} \|g - Hf\|^2 \right\} \]

- Hypothesis on \( f \): \( f \sim \mathcal{N}(0, \sigma^2_f (D^t D)^{-1}) \)
  \[ p(f) \propto \exp \left\{ -\frac{1}{2\sigma^2_f} \|Df\|^2 \right\} \]

- A posteriori:
  \[ p(f|g) \propto \exp \left\{ -\frac{1}{2\sigma^2_\epsilon} \|g - Hf\|^2 \frac{1}{2\sigma^2_f} \|Df\|^2 \right\} \]

- MAP:
  \[ \hat{f} = \arg \max_f \{ p(f|g) \} = \arg \min_f \{ J(f) \} \]
  with \( J(f) = \|g - Hf\|^2 + \lambda \|Df\|^2 \), \( \lambda = \frac{\sigma^2_\epsilon}{\sigma^2_f} \)

- Advantage: characterization of the solution

\[ f|g \sim \mathcal{N}(\hat{f}, \hat{P}) \] with \( \hat{f} = \hat{P} H^t g \), \( \hat{P} = (H^t H + \lambda D^t D)^{-1} \)
MAP estimation with other priors:

\[ \hat{f} = \arg \min_f \{ J(f) \} \quad \text{avec} \quad J(f) = \| g - Hf \|^2 + \lambda \Omega(f) \]

Separable priors:
- Gaussian prior:
  \[ p(f_j) \propto \exp\{-\alpha(f_j - m_j)^2\} \quad \Rightarrow \quad \Omega(f) = \alpha \sum_j (f_j - m_j)^2 \]
- Gamma prior:
  \[ p(f_j) \propto \left(\frac{f_j}{m_j}\right)^\alpha \exp\{-\frac{f_j}{m_j}\} \quad \Rightarrow \quad \Omega(f) = \alpha \sum_j \ln \frac{f_j}{m_j} + \frac{f_j}{m_j} , \]
- Beta prior:
  \[ p(f_j) \propto f_j^\alpha (1 - f_j)^\beta \quad \Rightarrow \quad \Omega(f) = \alpha \sum_j \ln f_j + \beta \sum_j \ln(1 - f_j) , \]
- Generalized gaussienne:
  \[ p(f_j) \propto \exp\{-\alpha|f_j - m_j|^p\}, \quad 1 < p < 2 \quad \Rightarrow \quad \Phi(f) = \alpha \sum_j |f_j - m_j|^p , \]

Markovian models:

\[ p(f_j|f_{-j}) \propto \exp\left\{-\alpha \sum_{i \in N_j} \phi(f_j, f_i)\right\} \quad \Rightarrow \quad \Phi(f) = \alpha \sum_j \sum_{i \in N_j} \phi(f_j, f_i) \]
MAP estimation with markovian priors:

\[ \hat{f} = \arg \min_f \{ J(f) \} \quad \text{avec} \quad J(f) = \| g - H f \|_2^2 + \lambda \Omega(f) \]

\[ \Omega(f) = \sum_j \phi(f_j - f_{j-1}) \]

with \( \phi(t) : \)

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

or

\[ \log(1 + t^2), \frac{t^2}{1 + t^2}, \arctan(t^2), \begin{cases} t^2 & |t| \leq T \\ T^2 & |t| > T \end{cases} \]
Which images I am looking for?
Which image I am looking for?

Gauss-Markov

Generalized GM

Piecewise Gaussian

Mixture of GM
Markovian prior models for images

\[ \Omega(f) = \sum_j \phi(f_j - f_{j-1}) \]

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

or equivalently:

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

\( q \) line process (contours)

- Mixture of Gaussians:

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

\( z \) region labels process.
Gauss-Markov-Potts prior models for images

\[ f(r) \quad z(r) \quad q(r) = 1 - \delta(z(r) - z(r')) \]

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

\[
p(f) = \sum P(z = k) \mathcal{N}(m_k, v_k) \quad \text{Mixture of Gaussians}
\]

Separable iid hidden variables:

\[
p(z) = \prod_r p(z(r))
\]

Markovian hidden variables:

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

\[
p(z) \propto \exp \left\{ -\gamma \sum_{r \in \mathcal{R}} \sum_{r' \in \mathcal{V}(r)} \delta(z(r) - z(r')) \right\}
\]
Four different cases

- $f|z$ iid, $z$ iid: Classical case of Mixture of Gaussians
- $f|z$ Markov, $z$ iid: (Markov composite) Mixture of Gauss-Markov
- $f|z$ iid, $z$ Markov: (Hidden Potts-Markov) Gauss-Potts
- $f|z$ Markov, $z$ Markov: (Gauss-Markov-Potts)
Case 1: \( f \mid z \text{ iid}, \quad z \text{ iid} \)

Independent Mixture of Independent Gaussians (IMIG):

\[
p(f(r) \mid z(r) = k) = \mathcal{N}(m_k, v_k), \quad \forall r \in \mathcal{R}
\]

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

\[
p(z) = \prod_r p(z(r) = k) = \prod_r \alpha_k = \prod_k \alpha_k^{n_k}
\]

Noting

\[
m_z(r) = m_k, \; v_z(r) = v_k, \; \alpha_z(r) = \alpha_k, \quad \forall r \in \mathcal{R}_k
\]

we have:

\[
p(f \mid z) = \prod_{r \in \mathcal{R}} \mathcal{N}(m_z(r), v_z(r))
\]

\[
p(z) = \prod_{r} \alpha_z(r) = \prod_k \alpha_k^{\sum_{r \in \mathcal{R}} \delta(z(r) - k)} = \prod_k \alpha_k^{n_k}
\]
Case 2: \( f | z \) Gauss-Markov, \( z \) iid

Independent Mixture of Gauss-Markov (IMGM):

\[
p(f(r)|z(r), z(r'), f(r'), r' \in \mathcal{V}(r)) = \mathcal{N}(\mu_z(r), \nu_z(r)), \forall r \in \mathcal{R}
\]

\[
\begin{align*}
\mu_z(r) &= \frac{1}{|\mathcal{V}(r)|} \sum_{r' \in \mathcal{V}(r)} \mu^*_z(r') \\
\mu^*_z(r') &= \delta(z(r') - z(r)) f(r') + (1 - \delta(z(r') - z(r))) m_z(r') \\
&= (1 - c(r')) f(r') + c(r') m_z(r')
\end{align*}
\]

\[
p(f|z) \propto \prod_r \mathcal{N}(\mu_z(r), \nu_z(r)) \propto \prod_k \alpha_k \mathcal{N}(m_k 1, \Sigma_k)
\]

\[
p(z) = \prod_r \nu_z(r) = \prod_k \alpha^k
\]

with \( 1_k = 1, \forall r \in \mathcal{R}_k \) and \( \Sigma_k \) a covariance matrix \((n_k \times n_k)\).
Case 3: $f|z$ Gauss iid, $z$ Potts

Gauss iid as in Case 1:

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

Potts-Markov

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

$$p(z) \propto \exp \left\{ \gamma \sum_{r \in R} \sum_{r' \in \mathcal{V}(r)} \delta(z(r) - z(r')) \right\}$$
Case 4: $f \mid z$ Gauss-Markov, $\not\sim$ Potts

Gauss-Markov as in Case 2:

\[ p(f(r) \mid z(r), z(r'), f(r'), r' \in \mathcal{V}(r)) = \mathcal{N}(\mu_z(r), \nu_z(r)), \forall r \in \mathcal{R} \]

\[
\begin{align*}
\mu_z(r) &= \frac{1}{|\mathcal{V}(r)|} \sum_{r' \in \mathcal{V}(r)} \mu_z^*(r') \\
\mu_z^*(r') &= \delta(z(r') - z(r)) \cdot f(r') + (1 - \delta(z(r') - z(r)) \cdot m_z(r')
\end{align*}
\]

\[ p(f \mid z) \propto \prod_r \mathcal{N}(\mu_z(r), \nu_z(r)) \propto \prod_k \alpha_k \mathcal{N}(m_k 1, \Sigma_k) \]

Potts-Markov as in Case 3:

\[ p(z) \propto \exp \left\{ \gamma \sum_{r \in \mathcal{R}} \sum_{r' \in \mathcal{V}(r)} \delta(z(r) - z(r')) \right\} \]
Full Bayesian approach

\[ g = H f + \epsilon \]

- Forward & errors model: \( p(g|f, \theta; M) \)
- Prior models \( p(f|\theta; M) \) and \( p(\theta|M) \)
- Bayes: \( p(f, \theta|g; M) = \frac{p(g|f,\theta; M)p(f|\theta; M)p(\theta|M)}{p(g|M)} \)
- Joint MAP:

\[
(\hat{f}, \hat{\theta}) = \arg \max_{(f, \theta)} \{ p(f, \theta|g; M) \}
\]

- Posterior means:

\[
\begin{align*}
\hat{f} & = \int f \, p(f, \theta|g; M) \, df \, d\theta \\
\hat{\theta} & = \int \theta \, p(f, \theta|g; M) \, df \, d\theta
\end{align*}
\]

- Evidence of the model:

\[
p(g|M) = \int \int p(g|f, \theta; M)p(f|\theta; M)p(\theta|M) \, df \, d\theta
\]
Bayesian Computation

- Direct computation and use of $p(f, \theta | g; \mathcal{M})$ is too complex
- Approximations:
  - Gauss-Laplace (Gaussian approximation)
  - Exploration (Sampling) using MCMC methods
  - Separable approximation (Variational techniques)
- Main idea:
  Approximate $p(f, \theta | g; \mathcal{M})$ by $q(f, \theta) = q_1(f) q_2(\theta)$
  - Choice of approximation criterion
  - Choice of appropriate families of probability laws for $q_1(f)$ and $q_2(\theta)$
Bayesian computation with Gauss-Markov-Potts prior models

\[
p(f, z, \theta | g) = \frac{p(g | f, \theta) \ p(f | z, \theta) \ p(z)}{p(g | \theta)}
\]

Approximations:

- \( f | z \text{ iid, } z \text{ iid} \):
  \[
  q(f, z, \theta | g) = q_1(f | z) \ q_2(z) \ q_3(\theta).
  \]

- \( f | z \text{ iid, } z \text{ Markov} \):
  \[
  q(f, z, \theta | g) = q_1(f | z) \ q_{2w}(z_w) \ q_{2b}(z_b) \ q_3(\theta).
  \]

- \( f | z \text{ Markov, } z \text{ iid} \):
  \[
  q(f, z, \theta | g) = q_{1w}(f_w | z) \ q_{1b}(f_b | z) \ q_2(z) \ q_3(\theta).
  \]

- \( f | z \text{ Markov, } z \text{ iid} \):
  \[
  q(f, z, \theta | g) = q_{1w}(f_w | z) \ q_{1b}(f_b | z) \ q_{2w}(z_w) \ q_{2b}(z_b) \ q_3(\theta)
  \]
Application of CT in NDT

Reconstruction from only 2 projections
Application in CT

\[ g | f \]
\[ g = H f + \epsilon \]
\[ g | f \sim \mathcal{N}(H f, \nu \epsilon I) \]
Gaussian

\[ f | z \]
 iid Gaussian
 or
 Gauss-Markov

\[ z \]
 iid
 or
 Potts

\[ c \]
\[ c(r) \in \{0, 1\} \]
 or
\[ 1 - \delta(z(r) - z(r')) \]
 binary
Proposed algorithm

\[ p(f, z, \theta|g) \propto p(g|f, z, \theta) p(f|z, \theta) p(\theta) \]

General scheme:

\[ \hat{f} \sim p(f|\hat{z}, \hat{\theta}, g) \rightarrow \hat{z} \sim p(z|\hat{f}, \hat{\theta}, g) \rightarrow \hat{\theta} \sim (\theta|\hat{f}, \hat{z}, g) \]

- Estimate \( f \) using \( p(f|\hat{z}, \hat{\theta}, g) \propto p(g|f, \theta) p(f|\hat{z}, \hat{\theta}) \)
  Needs optimisation of a quadratic criterion.

- Estimate \( z \) using \( p(z|\hat{f}, \hat{\theta}, g) \propto p(g|\hat{f}, \hat{z}, \hat{\theta}) p(z) \)
  Needs sampling of a Potts Markov field.

- Estimate \( \theta \) using
  \[ p(\theta|\hat{f}, \hat{z}, g) \propto p(g|\hat{f}, \sigma^2 \epsilon I) p(\hat{f}|\hat{z}, (m_k, v_k)) p(\theta) \]
  Conjugate priors \(\rightarrow\) analytical expressions.
Results

Original
Backprojection
Filtered BP
LS

Gauss-Markov+pos
GM+Line process
GM+Label process
Application in Microwave imaging

\[ g(\omega) = \int f(r) \exp\{-j(\omega \cdot r)\} \, dr + \epsilon(\omega) \]

\[ g(u, v) = \int f(x, y) \exp\{-j(ux + vy)\} \, dx \, dy + \epsilon(u, v) \]

\[ g = Hf + \epsilon \]
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)
Some references


Some references (suite)


Questions and Discussions

- Thanks for your attentions
- ...
- ...
- Questions?
- Discussions?
- ...
- ...
CT from two projections = Joint distribution from its marginals

\[ g_1(x) = \int f(x, y) \, dy \]
\[ g_2(y) = \int f(x, y) \, dx \]

Given the marginals \( g_1(x) \) and \( g_2(y) \), find the joint distribution \( f(x, y) \)

Infinite number of solutions

\[ f(x, y) = g_1(x) \, g_2(y) \, \Omega(G_1(x), G_2(y)) \]

\( \Omega(u, v) \) is a Copula:

\[ \Omega(u, 0) = 0, \quad \Omega(u, 1) = 1, \]
\[ \Omega(0, u) = 0, \quad \Omega(1, u) = 1 \]

\((x, y) \in [0, 1]^2, G_1(x) \) and \( G_2(y) \) are CDFs of \( g_1(x) \) and \( g_2(y) \)

Example: \( \Omega(u, v) = uv \)

Any link between geometrical structure of \( f \) and Copula functions?