Inverse Problems in Imaging systems and Computer Vision: From Regularization to Bayesian Inference

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

Journée inter-instituts d’Alembert/Farman
Imagerie de la matière et du vivant :
de l’instrument à l’algorithme : 18 mars 2010, ENS Cachan
Content

- Invers problems:
  Examples and general formulation
- Inversion methods:
  Analytical and Algebraic methods
- Deterministic methods:
  Least squares and Regularization
- Probabilistic methods:
  Bayesian inference
- Three main steps in Bayesian inference:
  - Forward models and likelihood
  - Prior models for images
  - Bayesian computation
- Applications (Computed Tomography)
- Conclusions
- Questions and Discussion
Inverse problems in imaging systems

Example 1: Measuring variation of temperature with a thermometer
- $f(t)$ variation of temperature over time
- $g(t)$ variation of length 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 4: Microwave, ultrasound or optical imaging
- $f(x, y)$ a section of a real 3D body
- $g_\phi(r)$ a line of diffracted wave measured at a given angle
Measuring variation of temperature with a thermometer

- $f(t)$ variation of temperature over time
- $g(t)$ variation of length of the liquid in thermometer

Forward model: Convolution

$$g(t) = \int f(t') h(t - t') \, dt' + \epsilon(t)$$

$h(t)$: 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 $f(t)$
Making an image with a camera, a microscope or a telescope

- \( f(x, y) \) real scene
- \( g(x, y) \) observed image

Forward model: Convolution

\[
g(x, y) = \int \int f(x', y') h(x - x', y - y') \, dx' \, dy' + \epsilon(x, y)
\]

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

Inverse problem: Image restoration

Given the forward model \( \mathcal{H} (\text{PSF } h(x, y))) \) 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

\[
g(x, y) = \int \int f(x', y') h(x - x', y - y') \, dx' \, dy' + \epsilon(x, y)
\]

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 radiograph \( g_\phi(r, z) \)

- **Forward model:**
  Line integrals or Radon Transform

\[
g_\phi(r) = \int_{L_{r,\phi}} f(x, y) \, dl + \epsilon_\phi(r)
= \int \int f(x, y) \delta(r - x \cos \phi - y \sin \phi) \, dx \, dy + \epsilon_\phi(r)
\]

- **Inverse problem:** Image reconstruction

Given the forward model \( \mathcal{H} \) (Radon Transform) and a set of data \( g_\phi(r_j), i = 1, \ldots, M, j = 1, \ldots, N \) find \( f(x, y) \)
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 \( f(x, y, z) \)
Microwave or ultrasound imaging

Mesaurs : scattered wave by the object $\phi_d(r_i)$
Unknown quantity : $f(r) = k_0^2(n^2(r) - 1)$
Intermediate quantity : $\phi(r)$ (total field)

TM : 2D Case :

$\phi_d(r_i) = \int\int_D G_m(r_i, r') \phi(r') f(r') \, dr', \ r_i \in S$

$\phi(r) = \phi_0(r) + \int\int_D G_o(r, r') \phi(r') f(r') \, dr', \ r \in D$

Born approximation $(\phi(r') \simeq \phi_0(r'))$ :

$\phi_d(r_i) = \int\int_D G_m(r_i, r') \phi_0(r') f(r') \, dr', \ r_i \in S$
General formulation of inverse problems

- General non linear inverse problems:
  \[ g(s) = [\mathcal{H} f(r)](s) + \epsilon(s), \quad r \in \mathcal{R}, \quad s \in \mathcal{S} \]

- Linear models:
  \[ g(s) = \int f(r) h(r, s) \, dr + \epsilon(s) \]

  If \( h(r, s) = h(r - s) \) \( \longrightarrow \) Convolution.

- Discrete data:
  \[ g(s_i) = \int h(s_i, r) f(r) \, dr + \epsilon(s_i), \quad i = 1, \ldots, m \]

- Inversion: Given the forward model \( \mathcal{H} \) and the data \( g = \{g(s_i), i = 1, \ldots, m\} \) estimate \( f(r) \)

- Well-posed and Ill-posed problems (Hadamard): existence, uniqueness and stability

- Need for prior information
Analytical methods in X ray Tomography

Radon:

\[ g(r, \phi) = \int_{L} f(x, y) \, dl \]

\[ f(x, y) = \left( \frac{1}{2\pi^2} \right) \int_{0}^{\pi} \int_{-\infty}^{+\infty} \frac{\partial}{\partial r} g(r, \phi) \frac{1}{(r - x \cos \phi - y \sin \phi)} \, dr \, d\phi \]
Filtered Backprojection method

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

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

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

**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**:

\[ g(r, \phi) \xrightarrow{\text{FT}} \mathcal{F}_1 \xrightarrow{\text{Filter}} |\Omega| \xrightarrow{\text{IFT}} g_1(r, \phi) \xrightarrow{\text{Backprojection}} \mathcal{B} \xrightarrow{\text{F}^{-1}} f(x, y) \]
Limitations: Limited angle or noisy data

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

\[ g(s_i) = \int h(s_i, r) f(r) \, dr + \epsilon(s_i), \quad i = 1, \ldots, M \]

- \( f(r) \) is assumed to be well approximated by

\[ f(r) \simeq \sum_{j=1}^{N} f_j b_j(r) \]

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

\[ g(s_i) = g_i \simeq \sum_{j=1}^{N} f_j \int h(s_i, r) b_j(r) \, dr, \quad i = 1, \ldots, M \]

\[ g = Hf + \epsilon \quad \text{with} \quad H_{ij} = \int h(s_i, r) b_j(r) \, dr \]

- \( H \) is huge dimensional
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:
    \[ \Delta(g, H(f)) = \|g - H(f)\|^2 = \sum_i |g_i - h_i(f)|^2 \]
  - \( L_p \):
    \[ \Delta(g, H(f)) = \|g - H(f)\|^p = \sum_i |g_i - h_i(f)|^p, \quad 1 < p < 2 \]
  - 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
    → Need for prior information

Functional space (Tikhonov):

\[ g = \mathcal{H}(f) + \epsilon \rightarrow J(f) = \| g - \mathcal{H}(f) \|^2_2 + \lambda \| Df \|^2_2 \]

Finite dimensional space (Philips & Towmey):
\[ g = \mathcal{H}(f) + \epsilon \]

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

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

- More general regularization:
  \[ J(f) = \Delta_1(g, \mathcal{H}(f)) + \lambda \Delta_2(f, f_0) \]

Limitations:
- Errors are considered implicitly white and Gaussian
- Limited prior information on the solution
- Lack of tools for the determination of the hyperparameters
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 contained in the data and coming from priors

Limitations:

- Practical implementation and cost of calculation
Bayesian estimation approach

\[ \mathcal{M} : \quad g = H f + \epsilon \]

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

- 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 regularization:**

Maximum A Posteriori (MAP):

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

with \( Q(g, Hf) = -\ln p(g|f) \) and \( \lambda \Omega(f) = -\ln p(f) \)
Case of linear models and Gaussian priors

\[ g = H f + \epsilon \]

- Forward model and hypothesis on the noise:
  \[ \epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2 I) \rightarrow p(g|f) \propto \exp \left\{ -\frac{1}{2\sigma_\epsilon^2} \| g - H f \|^2 \right\} \]

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

- A posteriori:
  \[ p(f|g) \propto \exp \left\{ - \left[ \frac{1}{2\sigma_\epsilon^2} \| g - H f \|^2 + \frac{1}{2\sigma_f^2} \| D f \|^2 \right] \right\} \]

- MAP:
  \[ \hat{f} = \arg \max_f \{ p(f|g) \} = \arg \min_f \{ J(f) \} \]
  with
  \[ J(f) = \| g - H f \|^2 + \lambda \| D f \|^2, \quad \lambda = \frac{\sigma_\epsilon^2}{\sigma_f^2} \]

- Advantage: characterization of the solution

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

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

Separable priors:

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

Markovian models:

\[
p(f_j|f) \propto \exp\left\{ -\alpha \sum_{i \in N_j} \phi(f_j, f_i) \right\} \rightarrow \Omega(f) = \alpha \sum_j \sum_{i \in N_j} \phi(f_j, f_i),
\]
Three main steps in the Bayesian approach

► Forward modeling and likelihood expression: (depends on application)
► 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

Piecewize Gaussian

Generalized GM

Mixture of GM
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?

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

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

Mixture of Gaussians

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

\[ g \mid f = H f + \epsilon \]
\[ g \mid f \sim \mathcal{N}(H f, \sigma^2 \epsilon I) \]
\[ f \mid z \text{ iid Gaussian} \]
\[ f \mid z \text{ or Gauss-Markov} \]
\[ z \text{ iid} \]
\[ z \text{ or Potts} \]

\[ c(r) \in \{0, 1\} \]
\[ 1 - \delta(z(r) - z(r')) \]
\text{binary}
Bayesian Computation

\[ p(f, z, \theta | g) \propto p(g | f, \nu_\epsilon) p(f | z, m, v) p(z | \gamma, \alpha) p(\theta) \]

\[ \theta = \{ \nu_\epsilon, (\alpha_k, m_k, v_k), k = 1, \ldots, K \} \quad p(\theta) \quad \text{Conjugate priors} \]

- Direct computation and use of \( p(f, z, \theta | g; M) \) 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
  \[ p(f, z, \theta | g; M) \]
  by
  \[ q(f, z, \theta) = q_1(f) q_2(z) q_3(\theta) \]

  - Choice of approximation criterion: \( KL(q : p) \)
  - Choice of appropriate families of probability laws for \( q_1(f), q_2(z) \) and \( q_3(\theta) \)
Proposed algorithm

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

General scheme (MCMC and Gibbs sampling):

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

Iterative algorithm:

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

Original

Backprojection

Filtered BP

LS

Gauss-Markov+pos

GM+Line process

GM+Label process
Application in Microwave imaging

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

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

\[ g = Hf + \epsilon \]
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 applications (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

\[ f_i(x, y) \rightarrow h(x, y) \rightarrow \epsilon_i(x, y) \rightarrow g_i(x, y) \]

Observation model: \[ g_i = H f_i + \epsilon_i, \quad i = 1, 2, 3 \]
Images fusion and joint segmentation

(with O. Féron)

\[
\begin{align*}
  g_i(r) &= f_i(r) + \epsilon_i(r) \\
  p(f_i(r)|z(r) = k) &= \mathcal{N}(m_{ik}, \sigma_{i k}^2) \\
  p(f|z) &= \prod_i p(f_i|z)
\end{align*}
\]
Data fusion in medical imaging  
(with O. Féron)

\[
\begin{align*}
    g_i(r) &= f_i(r) + \epsilon_i(r) \\
    p(f_i(r)|z(r) = k) &= \mathcal{N}(m_{ik}, \sigma_{ik}^2) \\
    p(f|z) &= \prod_i p(f_i|z)
\end{align*}
\]
Super-Resolution

(with F. Humblot)

Low Resolution images

⇒

High Resolution image
Joint segmentation of hyper-spectral images

(with N. Bali & A. Mohammadpour)

\[
\begin{align*}
    g_i(r) &= f_i(r) + \epsilon_i(r) \\
    p(f_i(r)|z(r) = k) &= \mathcal{N}(m_{ik}, \sigma_{ik}^2), \quad k = 1, \ldots, K \\
    p(f|z) &= \prod_i p(f_i|z) \\
    m_{ik} &\text{ follow a Markovian model along the index } i
\end{align*}
\]
Some references

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. Fkih-Salem (3D X ray Tomography)

My present PhD students:
- H. Ayasso (Optical Tomography, Variational Bayes)
- D. Pougaza (Tomography and Copula)

My colleagues in GPI (L2S) & collaborators in other institutes:
- B. Duchêne & A. Joisel (Inverse scattering and Microwave Imaging)
- N. Gac & A. Rabanal (GPU Implementation)
- Th. Rodet (Tomography)

Questions and Discussions