Bayesian inference for inverses problems in signal and image processing

Ali Mohammad-Djafari
Ph.D. Students & collaborators:
M. Ichir, O. Féron, P. Brault, A. Mohammadpour, Z. Chama
H. Snoussi, F. Humblot, S. Moussaoui
Laboratoire des Signaux et Systèmes
CNRS-SUPÉLEC-UPS
Supélec, Plateau de Moulon
91192 Gif-sur-Yvette, FRANCE.

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

- Inverses problems in image processing
- Multi sensor image processing problems
- Basics of Bayesian approach
- HMM modeling of images
- Examples of applications
  - Single channel image restoration
  - Fourier synthesis in optical imaging
  - Multi channel data fusion and joint segmentation
  - Video movie segmentation with motion estimation
  - Blind source (image) separation (BSS)
  - Hyperspectral image segmentation
- Bayesian image processing in wavelet domain
- Some results and conclusions
Inverses problems in image processing

- General non linear inverse problem:
  \[ g(\mathbf{r}) = [\mathcal{H}f(\mathbf{r}')]_{\mathbf{r}} + \epsilon(\mathbf{r}), \quad \mathbf{r} = (x, y) \in \mathcal{R}, \quad \mathbf{r}' = (x', y') \in \mathcal{R}' \]

- Linear model:
  \[ g(\mathbf{r}) = \int_{\mathcal{R}'} f(\mathbf{r}') h(\mathbf{r}, \mathbf{r}') \, d\mathbf{r}' + \epsilon(\mathbf{r}) \]

- Linear and translation invariante (convolution) model:
  \[ g(\mathbf{r}) = \int_{\mathcal{R}'} f(\mathbf{r}') h(\mathbf{r} - \mathbf{r}') \, d\mathbf{r}' + \epsilon(\mathbf{r}) = h(\mathbf{r}) \ast f(\mathbf{r}) + \epsilon(\mathbf{r}) \]

- Discretized version
  \[ g = Hf + \epsilon \]

where  \( g = \{ g(\mathbf{r}), \mathbf{r} \in \mathcal{R} \}, \epsilon = \{ \epsilon(\mathbf{r}), \mathbf{r} \in \mathcal{R} \} \) and  \( f = \{ f(\mathbf{r}'), \mathbf{r}' \in \mathcal{R}' \} \)
Single channel image restoration

\[ g(x, y) = f(x, y) * h(x, y) + \varepsilon(x, y) \]

Observation model: \[ g = Hf + \varepsilon \]
Fourier synthesis in optical imaging

\[ g(\omega) = \int f(r) \exp[-j\omega^t r] \, dr + \epsilon(\omega) \]

- Coherent imaging: \( G(g) = g \quad \rightarrow \quad g = Hf + \epsilon \)
- Non coherent imaging: \( G(g) = |g| \quad \rightarrow \quad g = H(f) + \epsilon \)

\( g = \{g(\omega), \omega \in \Omega\}, \quad \epsilon = \{\epsilon(\omega), \omega \in \Omega\} \quad \text{and} \quad f = \{f(r), r \in \mathcal{R}\} \)
Multi sensor image processing problems

- Disjoint multi sensors system:
  \[ g_i = H_i f_i + \epsilon_i, \quad i = 1, \cdots, M \]

- General multi input multi output (MIMO) system:
  \[ g_i = \sum_{j=1}^{N} H_{ij} f_j + \epsilon_i, \quad i = 1, \cdots, M \]

- General unknown mixing gain MIMO system:
  \[ g_i = \sum_{j=1}^{N} A_{ij} H_j f_j + \epsilon_i, \quad i = 1, \cdots, M \]

- Blind Sources Separation (BSS) problem:
  \[ g_i = \sum_{j=1}^{N} A_{ij} f_j + \epsilon_i, \quad i = 1, \cdots, M \]

where \( A = \{ A_{ij}, i = 1, \cdots, M, \ j = 1, \cdots, N \} \) is an unknown mixing matrix.
Multi-spectral image deconvolution

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

Observation model: \( g_i = Hf_i + \epsilon_i, \quad i = 1, 2, 3 \)
Image fusion and joint segmentation

$g_i(r) = f_i(r) + \epsilon_i(r), \quad i = 1, \cdots, M$

$\underline{g}(r) = \{g_i(r), \ i = 1, M\}$, \quad $\underline{g}_i = \{g_i(r), \ r \in \mathcal{R}\}$, \quad $\underline{g} = \{g_i(r), \ i = 1, M\}$

$\underline{g}(r) = \underline{f}(r) + \epsilon(r)$, \quad $\underline{g} = \underline{f} + \epsilon$
Blind image separation and joint segmentation

\[ g_i(r) = \sum_{j=1}^{N} A_{ij} f_j(r) + \epsilon_i(r) \]

\[ \hat{g}(r) = \{ g_i(r), \ i = 1, M \} \]

\[ \hat{g}(r) = A\hat{f}(r) + \epsilon(r), \]

\[ g = \{ g_i(r), \ i = 1, M \} \]

\[ g_i = \{ g_i(r), \ r \in \mathcal{R} \}, \]

\[ \underline{g} = A\underline{f} + \epsilon \]
Segmentation of hyperspectral images
Segmentation of hyperspectral images

\[ g_i(r) = f_i(r) + \epsilon_i(r), \quad i = 1, \cdots, M \]

\[ \underline{g}(r) = \{g_i(r), \quad i = 1, M\}, \quad g_i = \{g_i(r), \quad r \in R\}, \quad g = \{g_i(r), \quad i = 1, M\} \]

\[ \underline{g}(r) = \underline{f}(r) + \epsilon(r), \quad g = f + \epsilon \]
Blind image separation and joint segmentation

\[ g_i(r) = \sum_{j=1}^{N} A_{ij} f_j(r) + \epsilon_i(r) \]

\[ \underline{g}(r) = \{g_i(r), \ i = 1, M\} \]

\[ \underline{g}(r) = A\underline{f}(r) + \epsilon(r), \]

\[ g = \{g_i(r), \ i = 1, M\} \]

\[ g_i = \{g_i(r), \ r \in \mathcal{R}\}, \]

\[ \underline{g} = A\underline{f} + \epsilon \]
Bayesian estimation approach

\[ g = Af + \epsilon \]

- Forward model and prior knowledge on the noise \[ \rightarrow p(g|f) \]
- Prior knowledge on \( f \) \[ \rightarrow p(f) \]
- Bayes rule:
  \[ p(f|g) = \frac{p(g|f) \ p(f)}{\int p(g|f) \ p(f) \ df} \]
- Infer on \( f \) via \( p(f|g) \):
  - MAP estimator:
    \[ \hat{f} = \arg \max_f \{ p(f|g) \} = \arg \min_f \{ - \ln p(f|g) \} \]
    \[ \hat{f} = \int f \ p(f|g) \ df \] posterior mean
Bayesian blind sources separation (BBSS)

\[ g = Af + \epsilon \]

- Likelihood \( p(g|f, A, \theta_1) \)
- Priors \( p(f|\theta_2), p(A|\theta_3) \) and \( p(\theta) \) hyperparameters \( \theta = \{\theta_1, \theta_2, \theta_3\} \)
- Bayes rule
  \[ p(f, A, \theta|g) \propto p(g|f, A, \theta_1) p(f|\theta_2) p(A|\theta_3) p(\theta) \]
- Generalized a posteriori EM-ML estimator
  \[ (\hat{A}, \hat{\theta}) = \arg\max_{(A, \theta)} \{p(A, \theta|g)\} \quad \rightarrow \quad \hat{f} = \arg\max_{f} \{p(f|g, \hat{A}, \hat{\theta})\} \]
- MSE estimator which corresponds to the posterior mean using MCMC
  \[ \{\hat{f}, \hat{A}, \hat{\theta}\} = \int \{f, A, \theta\} p(f, A, \theta|g) \, d\{f, A, \theta\}. \]
HMM modeling of images

What they share?
Borders & Regions
Segmentation
Hidden variable $z(r)$

$z(r) = k, \ k = 1, \cdots, K$

$\mathcal{R}_k = \{ r : z(r) = k \}, \ \mathcal{R} = \bigcup_k \mathcal{R}_k$

Homogeneity in regions

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

$R_{jk} = \{ r : z_j(r) = k \}, \quad \mathcal{R}_j = \bigcup_k R_{jk}$

$f_{jk} = \{ f_j(r) : r \in R_{jk} \}, \quad f_j = \bigcup_k f_{jk}$

$p(f_j(r), r \in \mathcal{R}) = \prod_{k=1}^{K_j} p(f_j(r), r \in R_{jk}) = \prod_{k=1}^{K_j} \mathcal{N}(m_{jk}, \sigma_{jk}^2)$
Main hypothesis:

- Pixels values in different regions of an image are independent.

- For pixels values in a given region of an image, two possibilities:
  - i.i.d.: \( p(f_j(r)|z_j(r) = k) = \mathcal{N}(m_{jk}, \sigma_{jk}^2) \)
  \[ p(f_j(r), r \in R_{jk}) = \mathcal{N}(m_{jk}1, \sigma_{jk}^2 I) \]
  - Markovien: \( p(f_j(r), r \in R_{jk}) = \mathcal{N}(m_{jk}1, \Sigma_{jk}) \)

- For pixels values in different images but in a given common region two possibilities:
  - i.i.d.: \((f_j(r)|z(r) = k)\) independent of \((f_i(r)|z(r) = k)\), \(i \neq j\)
  - Markovien: \( p(f_j(r)|z(r) = k, \underline{f}(r)) = p(f_j(r)|z(r) = k, f_{j-1}(r)) \)
Modeling the labels

\[ p(f_j(r)|z_j(r) = k) = \mathcal{N}(m_{jk}, \sigma^2_{jk}) \rightarrow p(f_j(r)) = \sum_k P(z_j(r) = k) \mathcal{N}(m_{jk}, \sigma^2_{jk}) \]

- **Independent Gaussian Mixture model (IGM)**, where \( z_j = \{ z_j(r), r \in \mathcal{R} \} \) are assumed to be independent and

\[ P(z_j(r) = k) = p_k, \quad \text{with} \quad \sum_k p_k = 1 \quad \text{and} \quad p(z_j) = \prod_k p_k \]

- **Contextual Gaussian Mixture model (CGM)**: \( z_j \) Markovien

\[ p(z_j) \propto \exp \left[ \alpha \sum_{r \in \mathcal{R}} \sum_{s \in \mathcal{V}(r)} \delta(z_j(r) - z_j(s)) \right] \]

which is the Potts Markov random field (PMRF).

The parameter \( \alpha \) controls the mean value of the regions’ sizes.
Expressions of likelihood, prior and posterior laws

\[ g_i = \sum_{j=1}^{N} A_{i,j} f_j + \epsilon_i, \quad i = 1, \cdots, M \quad \rightarrow \quad g = A \underline{f} + \epsilon \]

- Likelihood:

\[ p(g|A, f, \theta_1) = \prod_{i=1}^{M} p(g_i|A, \underline{f}, \Sigma_{\epsilon_i}) = \prod_{i=1}^{M} \mathcal{N}(A \underline{f}, \Sigma_{\epsilon_i}) \]

\[ \Sigma_{\epsilon_i} = \sigma_{\epsilon_i}^2 I, \quad \rightarrow \quad \theta_1 = \{\sigma_{\epsilon_i}^2, i = 1, \cdots, M\} \]

- HMM for the images:

\[ p(f|z, \theta_2) = \prod_{j=1}^{N} p(f_j|z_j, m_j, \Sigma_j) \]

where \( z = \{z_j, j = 1, \cdots, N\} \) and where we assumed that \( f_j|z_j \) are independent. \( \rightarrow \theta_2 = \{(m_{ik}, \sigma_{jk}^2), j = 1, \cdots, N\} \)
• PMRF for the labels:

\[ p(z) \propto \prod_{j=1}^{N} \exp \left[ \alpha \sum_{r \in R} \sum_{s \in \mathcal{V}(r)} \delta(z_j(r) - z_j(s)) \right] \]

• Conjugate priors for the hyperparameters \( \theta = (\theta_1, \theta_2) \):

\[ \theta = \{\{\sigma_{\epsilon_i}^2, i = 1, \cdots, M\}, \{(m_{ik}, \sigma_{j_k}^2), j = 1, \cdots, M, k = 1, \cdots, K\}\} \]

\[ p(m_{jk}) = \mathcal{N}(m_{jk0}, \sigma_{jk0}^2) \]

\[ p(\sigma_{jk}^2) = \mathcal{IG}(\alpha_{j0}, \beta_{j0}) \]

\[ p(\Sigma_{jk}) = \mathcal{IW}(\alpha_{j0}, \Lambda_{j0}) \]

\[ p(\sigma_{\epsilon_i}) = \mathcal{IG}(\alpha_{i0}, \beta_{i0}) \]

• Joint posterior law of \( f, z \) and \( \theta \)

\[ p(f, z, \theta | g) \propto p(g | f, \theta_1) p(f | z, \theta_2) p(z | \theta_2) p(\theta) \]
General MCMC sampling scheme

\[ p(\mathbf{f}, \mathbf{z}, \theta | \mathbf{g}) \propto p(\mathbf{g} | \mathbf{f}, \theta_1) \ p(\mathbf{f} | \mathbf{z}, \theta_2) \ p(\mathbf{z} | \theta_2) \ p(\theta) \]

Gibbs sampling:

- Generate samples \((\mathbf{f}, \mathbf{z}, \theta)^{(1)}\), \(\cdots\), \((\mathbf{f}, \mathbf{z}, \theta)^{(N)}\) using
  - \( \mathbf{f} \sim p(\mathbf{f} | \mathbf{g}, \mathbf{z}, \theta) \),
  - \( \mathbf{z} \sim p(\mathbf{z} | \mathbf{g}, \mathbf{f}, \theta) \),
  - \( \theta \sim p(\theta | \mathbf{g}, \mathbf{f}, \mathbf{z}) \),

- Compute any statistics such as mean, median, variance, ...

Difficulties:

- No mixture, No convolution: \( g_i = f_i + \varepsilon_i, \ i = 1, \cdots, M \)
- Mixture but No convolution: \( g_i = \sum_{j=1}^{N} A_{ij} f_j + \varepsilon_i, \ i = 1, \cdots, M \)
- Convolution but No mixture: \( g_i = H_i f_i + \varepsilon_i, \ i = 1, \cdots, M \)
- Mixture and Convolution: \( g_i = \sum_{j=1}^{N} A_{ij} H_{ij} f_j + \varepsilon_i, \ i = 1, \cdots, M \)
Examples of applications

- No mixture, No convolution applications:
  - Multi channel image fusion and joint segmentation
    \[ g_i = f_i + \epsilon_i, \quad i = 1, \ldots, M, \quad f_i | z \text{ independent} \]
  - Hyperspectral image segmentation
    \[ g_i = f_i + \epsilon_i, \quad i = 1, \ldots, M, \quad f_i | z \text{ dependent} \]
  - Video movie segmentation with motion estimation
    \[ g_i = f_i + \epsilon_i, \quad i = 1, \ldots, M, \quad f_i | z_i \text{ independent} \]

- Mixture, No convolution applications:
  - Blind source (image) separation (BSS) and joint segmentation
    \[ g_i = \sum_{j=1}^{N} A_{ij} f_j + \epsilon_i, \quad i = 1, \ldots, M, \quad f_i | z_i \text{ independent} \]

- Convolution but No mixture applications:
  - Fourier synthesis in optical imaging
    \[ g = H(f) + \epsilon, \quad f | z \]
  - Single channel image restoration
    \[ g = Hf + \epsilon, \quad f | z \]
Images fusion and joint segmentation

(Olivier FÉRON)

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

(Adel MOHammAdPOUR)

\[
\begin{align*}
g_i(r) &= f_i(r) + \epsilon_i(r), \quad r \in \mathcal{R}, \quad \text{or} \quad g_i = f_i + \epsilon_i, \quad i = 1, \ldots, M \\
p(f_i(r)|z(r) = k) &= \mathcal{N}(m_{ik}, \sigma^2_{ik}), \quad k = 1, \ldots, K \\
p(f|z) &= \prod_i p(f_i|z) \\
m_{ik} \quad \text{follow a Markovian model along the index} \quad i
\end{align*}
\]
Segmentation of a video sequence of images

(Patrice BRAULT)

\[
\begin{aligned}
g_i(r) &= f_i(r) + \epsilon_i(r), \quad r \in \mathcal{R}, \quad \text{or} \quad g_i = f_i + \epsilon_i, \quad \text{or} \quad g = f + \epsilon \\
p(f_i(r)|z_i(r) = k) &= \mathcal{N}(m_{ik}, \sigma^2_{ik}), \quad k = 1, \cdots, K \\
p(f|z) &= \prod_i p(f_i|z_i) \\
z_i(r) \quad \text{follow a Markovian model along the index} \quad i
\end{aligned}
\]
Blind image separation and joint segmentation

\[
g_i(r) = \sum_{j=1}^{N} A_{ij} f_j(r) + \epsilon_i(r) \quad \rightarrow \quad g(r) = Af(r) + \epsilon(r) \quad \rightarrow \quad g = Af + \epsilon
\]

\[
p(g|f, A, \Sigma \epsilon) = \prod_{r \in \mathcal{R}} p(g(r)|f(r), A) = \prod_{r \in \mathcal{R}} \mathcal{N}(Af(r), \Sigma \epsilon)
\]

\[
p(f_j(r)|z_j(r) = k) = \mathcal{N}(m_{jk}, \sigma_{jk}^2),
\]

\[
p(f|z) = \prod_{r \in \mathcal{R}} \prod_j p(f_j(r)|z_j(r))
\]

\[
p(A_{ij}) = \mathcal{N}(A_{0ij}, \sigma_{0ij}^2) \quad \text{or} \quad p(\text{vect}(A)) = \mathcal{N}(\text{vect}(A_0), \sigma_{0ij}^2 I)
\]

\[
p(f|g, z, \theta, A) = \sum_{r \in \mathcal{R}} \mathcal{N}(\hat{f}(r), \hat{\Sigma}(r)) \quad \text{with}
\]

\[
\hat{\Sigma}(r) = (A^t \Sigma \epsilon^{-1} A + \Sigma z^{-1}(r))^{-1}
\]

\[
\hat{f}(r) = \hat{\Sigma}(r) (A^t \Sigma \epsilon^{-1} g(r) + \Sigma z(r)^{-1} m_z(r))
\]

\[
p(z(r)|g(r), \theta, A) \propto (p(g(r)|z(r), \theta, A)) \quad p(z(r)) \quad \text{with}
\]

\[
p(g(r)|z(r), \theta) = \mathcal{N}(Am_z(r), A \Sigma z(r) A^t + \Sigma \epsilon)
\]

\[
p(A, \Sigma \epsilon|f, g, \theta) = \mathcal{N}(A; A_p, \Sigma_p) \mathcal{W}(\Sigma \epsilon^{-1}; \nu_p, \Sigma \epsilon_p)
\]
Single channel image restoration

\[ g(r') = \int h(r' - r) f(r) \, dr + \epsilon(r') \rightarrow g = H f + \epsilon \]

Fourier synthesis inverse problem

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

\[ p(\epsilon) = \mathcal{N}(0, \Sigma_\epsilon) \rightarrow p(g|f) = \mathcal{N}(H f, \Sigma_\epsilon) \quad \text{with} \quad \Sigma_\epsilon = \sigma_\epsilon^2 I \]

\[ p(f(r)|z(r) = k) = \mathcal{N}(m_k, \sigma_k^2), \quad k = 1, \ldots, K \]

\[ p(f|z, \theta, g) = \mathcal{N}(\hat{f}, \hat{\Sigma}) \quad \text{with} \quad \hat{\Sigma} = (H^t \Sigma_\epsilon^{-1} H + \Sigma_z^{-1})^{-1} \quad \text{and} \quad \hat{f} = \hat{\Sigma} (H^t \Sigma_\epsilon^{-1} g + \Sigma_z^{-1} m_z) \]

Compute \( \hat{f} = \arg \max_f \{ p(f|z, \theta, g) \} = \arg \min_f \{ J(f) \} \quad \text{with} \quad J(f) = \frac{1}{\sigma_\epsilon^2} \| g - H f \|^2 + \sum_k \frac{\| f_k - m_k \|^2}{\sigma_k^2} \]

\[ p(z|g, \theta) \propto p(g|z, \theta) \, p(z) \quad \text{with} \]

\[ p(g|z, \theta) = \mathcal{N}(H m_z, \Sigma g) \quad \text{with} \quad \Sigma g = H \Sigma_z H^t + \Sigma_\epsilon \]

Use \( p(z|g, f, \theta) \propto p(g|f, z, \theta) \, p(z) \)
Simulation results

a) object, b) exact known support, c) support of the data, d) measured data, 
e) and f) Results when phase is measured: e) IFT and f) proposed method,  
g) and h) Results when the phase is not measured but we know the support of  
the object: g) by Gerchberg-Saxton h) by the proposed method.
Wavelet domain Bayesian image processing

\[ g(r) = Af(r) + \epsilon(r) \longrightarrow \text{WT} \longrightarrow g^j(r) = Af^j(r) + \epsilon^j(r) \]

- multi-resolution computation
- Wavelet coefficients can be classified and segmented in only \( K = 2 \) classes
Conclusion and works in progress

Bayesian approach & HMM are appropriate tools for many image processing problems

- **H. Snoussi**: BSS in 1D and 2D either in pixel domain or Fourier transform domain
- **M. Ichir**: BSS with mixture of Gamma and BSS in wavelet domain
- **S. Moussaoui**: BSS for non-negative sources with application in spectrometry
- **O. Féron**: Data and image fusion, Fourier synthesis and inverse problems in microwave imaging
- **P. Brault**: Segmentation of images sequences either directly or in wavelet domain
- **A. Mohammadpour**: Segmentation of hyper-spectral images,
- **Z. Chama**: Image recovery from the Fourier phase (Fourier Synthesis)
- **F. Humblot**: Obtaining a super-resolution image from a set of lower resolution images