Statistical methods 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, N. Bali, H. Akbari

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

Inverse problems

$\mathcal{H}$	(	$\boldsymbol{g}$,	$\boldsymbol{f}$,	$\boldsymbol{z}$,	$\boldsymbol{\epsilon}$	)	=	0
model		measurement (Data)	unknown quantities	other unknowns	errors and noise			

Explicit relation: $\boldsymbol{g} = \mathcal{H}(\boldsymbol{f}, \boldsymbol{z}, \boldsymbol{\epsilon})$

Additive output error model: $\boldsymbol{g} = \mathcal{H}(\boldsymbol{f}, \boldsymbol{z}) + \boldsymbol{\epsilon}$

Relation between $\boldsymbol{f}$ and $\boldsymbol{z}$:
$$\left\{ \begin{array}{l} \boldsymbol{g} = \mathcal{H}_1(\boldsymbol{f}, \boldsymbol{z}) + \boldsymbol{\epsilon} \\ \mathcal{H}_2(\boldsymbol{f}, \boldsymbol{z}) = 0 \end{array} \right.$$

Non linear+ additive errors: $g_i = h_i(\boldsymbol{f}) + \epsilon_i$ or $\boldsymbol{g} = \boldsymbol{H}(\boldsymbol{f}) + \boldsymbol{\epsilon}$

Linear + additive errors: $h_i(\boldsymbol{f}) = \sum_j h_{ij} f_j \longrightarrow \boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\epsilon}$

Inverses problems in image processing

- General non linear inverse problem:

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

- Linear model:

$$g(\boldsymbol{r}) = \int_{\mathcal{R}'} f(\boldsymbol{r}') h(\boldsymbol{r}, \boldsymbol{r}') \,\mathrm{d}\boldsymbol{r}' + \epsilon(\boldsymbol{r})$$

- Linear and translation invariante (convolution) model:

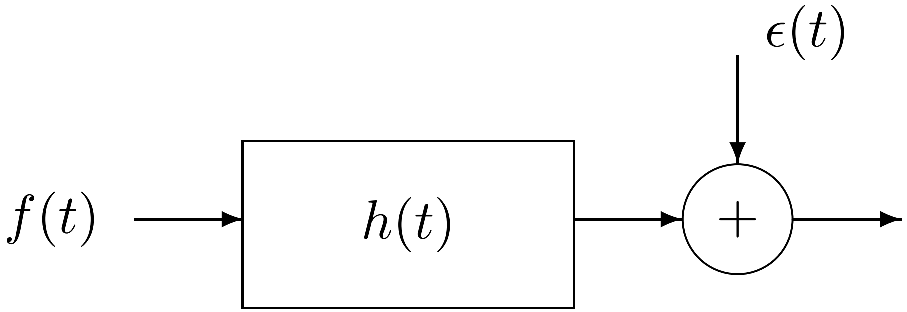
$$g(\boldsymbol{r}) = \int_{\mathcal{R}'} f(\boldsymbol{r}') h(\boldsymbol{r} - \boldsymbol{r}') \,\mathrm{d}\boldsymbol{r}' + \epsilon(\boldsymbol{r}) = h(\boldsymbol{r}) * f(\boldsymbol{r}) + \epsilon(\boldsymbol{r})$$

- Discretized version

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

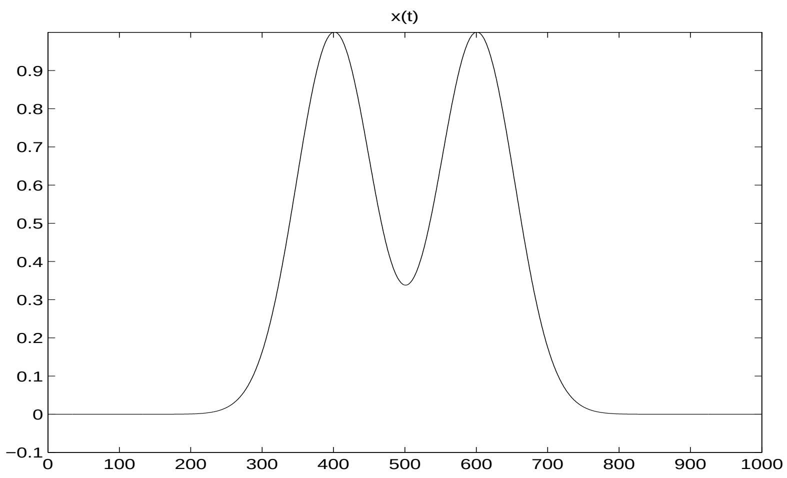
where $\boldsymbol{g} = \{g(\boldsymbol{r}),\ \boldsymbol{r} \in \mathcal{R}\}$, $\boldsymbol{\epsilon} = \{\epsilon(\boldsymbol{r}),\ \boldsymbol{r} \in \mathcal{R}\}$ and $\boldsymbol{f} = \{f(\boldsymbol{r}'),\ \boldsymbol{r}' \in \mathcal{R}'\}$

Deconvolution

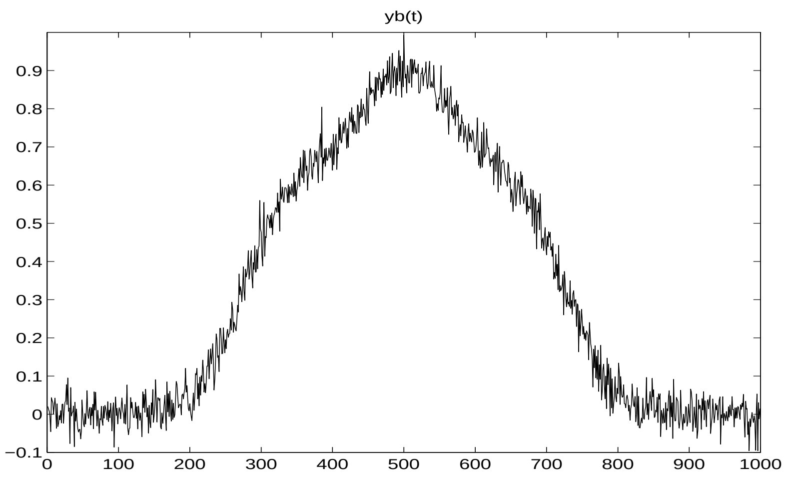


$$g(t) = h(t) * f(t) + \epsilon(t)$$

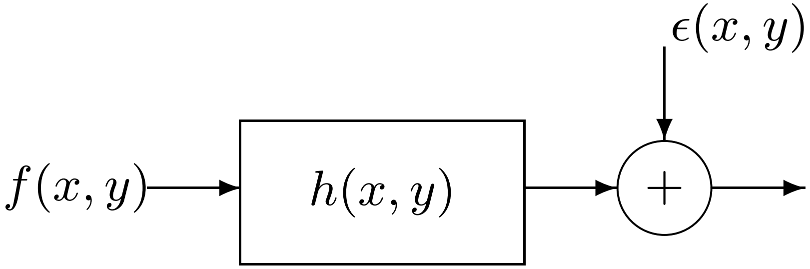
Observation model: $\boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\epsilon}$



$\overset{?}{\Longleftarrow}$



Single channel image restoration



$g(x,y) = f(x,y) * h(x,y) + \epsilon(x,y)$

Observation model : $\boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\epsilon}$

?

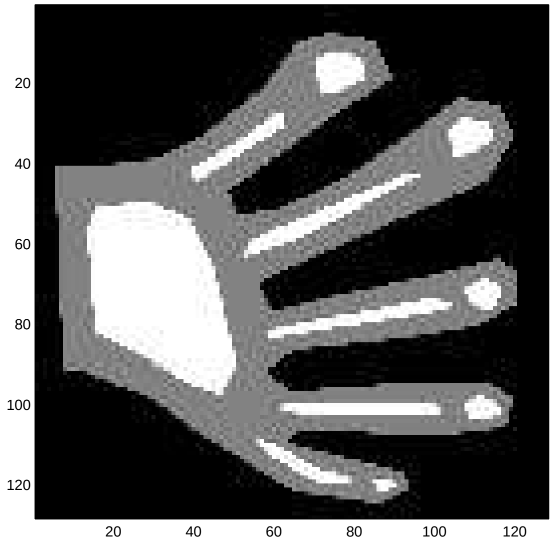
$\Longleftarrow$

Fourier synthesis in optical imaging

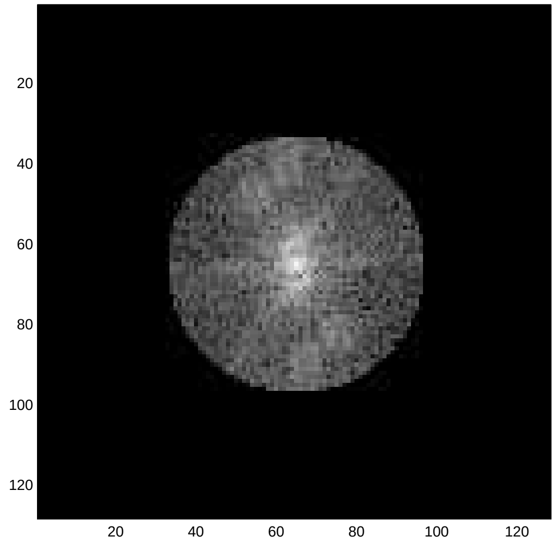
$$g(\boldsymbol{\omega}) = \int f(\boldsymbol{r}) \exp\left[-j\boldsymbol{\omega}^t \boldsymbol{r}\right] \, \mathrm{d}\boldsymbol{r} + \epsilon(\boldsymbol{\omega})$$

- Coherent imaging: $\mathcal{G}(g) = g \quad \longrightarrow \quad \boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\epsilon}$
- Non coherent imaging: $\mathcal{G}(g) = |g| \quad \longrightarrow \quad \boldsymbol{g} = \boldsymbol{H}(\boldsymbol{f}) + \boldsymbol{\epsilon}$

$$\boldsymbol{g} = \{g(\boldsymbol{\omega}),\ \boldsymbol{\omega} \in \Omega\}, \quad \boldsymbol{\epsilon} = \{\epsilon(\boldsymbol{\omega}),\ \boldsymbol{\omega} \in \Omega\} \quad \text{and} \quad \boldsymbol{f} = \{f(\boldsymbol{r}),\ \boldsymbol{r} \in \mathcal{R}\}$$



$$? \qquad \Longleftarrow$$



Data matching

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

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

- Examples:
  - LS $\qquad \Delta(\boldsymbol{g}, \boldsymbol{H}(\boldsymbol{f})) = \|\boldsymbol{g} - \boldsymbol{H}(\boldsymbol{f})\|^2 = \sum_i |g_i - h_i(\boldsymbol{f})|^2$
  - L_p $\qquad \Delta(\boldsymbol{g}, \boldsymbol{H}(\boldsymbol{f})) = \|\boldsymbol{g} - \boldsymbol{H}(\boldsymbol{f})\|^p = \sum_i |g_i - h_i(\boldsymbol{f})|^p, \quad 1 < p < 2$
  - KL $\qquad \Delta(\boldsymbol{g}, \boldsymbol{H}(\boldsymbol{f})) = \sum_i g_i \ln \frac{g_i}{h_i(\boldsymbol{f})}$
- In general, does not give satisfactory results for inverse problems.

Regularization theory

Inverse problems = Ill posed problems

$\longrightarrow$ Need of prior information

- 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:

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

or

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

Limitations:

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

Probabilistic methods

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

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

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

Maximum Likelihood (ML)

- Data observation Model: $g_i = h_i(\boldsymbol{f}) + \epsilon_i$ or $\boldsymbol{g} = \boldsymbol{H}(\boldsymbol{f}) + \boldsymbol{\epsilon}$
- Noise probability law: $p_\epsilon(\boldsymbol{\epsilon}) \longrightarrow p(\boldsymbol{g}|\boldsymbol{f}) \longrightarrow l(\boldsymbol{f}) \stackrel{\triangle}{=} p(\boldsymbol{g}|\boldsymbol{f})$
- ML estimate

$$\widehat{\boldsymbol{f}} = \arg\max_{\boldsymbol{f}} \{l(\boldsymbol{f})\} = \arg\min_{\boldsymbol{f}} \{-\ln l(\boldsymbol{f})\}$$

- $l(\boldsymbol{f})$ is not a probability distribution
- $l(\boldsymbol{f})$ can be used to compare different models via the

$$\text{likelihood ratio} \quad \frac{l(\boldsymbol{f}_1)}{l(\boldsymbol{f}_2} > 1$$

- Case of linear model and Gaussian law: $\longrightarrow$ LS
- Rarely gives satisfactory solutions for inverse problems with limited number of data
- Penalized ML $\widehat{\boldsymbol{f}} = \arg\min_{\boldsymbol{f}} \{-\ln p(\boldsymbol{g}|\boldsymbol{f}) + \Phi(\boldsymbol{f})\}$

Minimum Inaccuracy (MI)

- Empirical density (histogram) $\quad \rho(\boldsymbol{g}) \stackrel{\triangle}{=} \frac{1}{N} \sum_i \delta(g - g_i)$
- Inaccuracy of $\rho(\boldsymbol{g})$ with respect to $p(\boldsymbol{g}|\boldsymbol{f})$

$$I[\rho, p] \stackrel{\triangle}{=} - \int \rho(\boldsymbol{g}) \ln p(\boldsymbol{g}|\boldsymbol{f}) \,\mathrm{d}\boldsymbol{g}$$

- Minimum Inaccuracy (MI) estimate:

$$\widehat{\boldsymbol{f}} = \arg\min_{\boldsymbol{f}} \left\{ I\left(\rho(\boldsymbol{g}), p(\boldsymbol{g}|\boldsymbol{f})\right) \right\}$$

- In case of i.i.d. data:

$$\begin{aligned} p(\boldsymbol{g}|\boldsymbol{f}) &= \prod_i p(g_i|\boldsymbol{f}) = \exp\left[N \frac{1}{N} \sum_i \ln p(g_i|\boldsymbol{f}) \right] \\ &= \exp\left[N \int \rho(\boldsymbol{g}) \ln p(\boldsymbol{g}|\boldsymbol{f}) \,\mathrm{d}\boldsymbol{g} \right] = \exp\left[-N I\left[\rho(\boldsymbol{g}), p(\boldsymbol{g}|\boldsymbol{f})\right] \right] \end{aligned}$$

$\longrightarrow \qquad$ MI = ML

Probability Distribution Matching (PDM)

- Match probability distributions in place of matching the data.
- Mismatch measure: Kullback-Leibler (KL) mismatch function

$$K[\rho, p] \stackrel{\triangle}{=} \int \rho(\boldsymbol{g}) \ln \frac{\rho(\boldsymbol{g})}{p(\boldsymbol{g}|\boldsymbol{f})} \, \mathrm{d}\boldsymbol{g}$$

- MKL estimate $\quad \widehat{\boldsymbol{f}} = \arg\min_{\boldsymbol{f}} \{K[\rho(\boldsymbol{g}), p(\boldsymbol{g}|\boldsymbol{f})]\}$
- Note that

$$\begin{aligned} K[\rho(\boldsymbol{g}), p(\boldsymbol{g}|\boldsymbol{f})] &= \int \rho(\boldsymbol{g}) \ln \frac{\rho(\boldsymbol{g})}{p(\boldsymbol{g}|\boldsymbol{f})} \, \mathrm{d}\boldsymbol{g} \\ &= -\int \rho(\boldsymbol{g}) \ln p(\boldsymbol{g}|\boldsymbol{f}) \, \mathrm{d}\boldsymbol{g} + \int \rho(\boldsymbol{g}) \ln \rho(\boldsymbol{g}) \, \mathrm{d}\boldsymbol{g} \\ &= I[\rho(\boldsymbol{g}), p(\boldsymbol{g}|\boldsymbol{f})] - H[\rho(\boldsymbol{g})], \end{aligned}$$

where $H[\rho(\boldsymbol{g})]$ is the entropy of $\rho(\boldsymbol{g})$ & does not depend on $\boldsymbol{f}$

$\longrightarrow \qquad$ PDM = MI

Maximum entropy (ME)

Data model: $\boldsymbol{g} = \boldsymbol{H}(\langle \boldsymbol{f} \rangle)$ where $\langle \boldsymbol{f} \rangle = \int \boldsymbol{f}\, p(\boldsymbol{f})\, \mathrm{d}\boldsymbol{f}$

$$\text{maximize} \quad -\int p(\boldsymbol{f}) \ln \frac{p(\boldsymbol{f})}{p_0(\boldsymbol{f})}\, \mathrm{d}\boldsymbol{f} \quad \text{s.t.} \quad \boldsymbol{g} = \boldsymbol{H}(\langle \boldsymbol{f} \rangle)$$

Solution:

$$p(\boldsymbol{f}) = \frac{1}{Z(\boldsymbol{\lambda})} p_0(\boldsymbol{f}) \exp\left[\boldsymbol{\lambda}^t \boldsymbol{H}(\boldsymbol{f})\right] = p_0(\boldsymbol{f}) \exp\left[\boldsymbol{\lambda}^t \boldsymbol{H}(\boldsymbol{f}) - \ln Z(\boldsymbol{\lambda})\right]$$

$$Z(\boldsymbol{\lambda}) = \int p_0(\boldsymbol{f}) \exp\left[\boldsymbol{\lambda}^t \boldsymbol{H}(\boldsymbol{f})\right]\, \mathrm{d}\boldsymbol{f}$$

$$\frac{\partial \ln Z(\boldsymbol{\lambda})}{\partial \lambda_i} = g_i, \quad i = 1, \cdots, M \longrightarrow \boldsymbol{\lambda} \longrightarrow p \longrightarrow \widehat{\boldsymbol{f}}(\boldsymbol{\lambda}) = \int \boldsymbol{f}\, p(\boldsymbol{f})\, \mathrm{d}\boldsymbol{f}$$

In case of linear models $\boldsymbol{H}(\boldsymbol{f}) = \boldsymbol{H}\boldsymbol{f}$ and noting by:

$$\boldsymbol{s} = \boldsymbol{H}^t\boldsymbol{\lambda}, \quad \Omega(\boldsymbol{s}) = \ln Z(\boldsymbol{s}), \quad D(\boldsymbol{\lambda}) = \boldsymbol{\lambda}^t\boldsymbol{g} - G(\boldsymbol{H}^t\boldsymbol{\lambda})$$

and $\quad \boldsymbol{m} = \int \boldsymbol{f}\, p_0(\boldsymbol{f})\,\mathrm{d}\boldsymbol{f}, \quad$ we obtain

$$\widehat{\boldsymbol{\lambda}} = \arg\min_{\boldsymbol{\lambda}} \{D(\boldsymbol{\lambda})\} \qquad \text{Dual criterion}$$

$$\widehat{\boldsymbol{f}} = \arg\min_{\boldsymbol{f}} \{\Phi(\boldsymbol{f},\boldsymbol{m})\} \text{ s.t. } \boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} \qquad \text{Primal criterion}$$

- Functions $D(\boldsymbol{\lambda})$ and $\Phi(\boldsymbol{f},\boldsymbol{m})$ depend on $p_0(\boldsymbol{f})$
- $\Phi(\boldsymbol{f},\boldsymbol{m})$ is a distance measure between $\boldsymbol{f}$ and $\boldsymbol{m}$
- $\exists$ analytical expressions for $D(\boldsymbol{\lambda})$ and $\Phi(\boldsymbol{f},\boldsymbol{m})$ for special cases of $p_0(\boldsymbol{f})$, in particular when $p_0(\boldsymbol{f})$ is separable.

- **Some special cases:**

Gaussian	$\Phi(\boldsymbol{f}, \boldsymbol{m}) = \sum_j (f_j - m_j)^2$	Quadratic
Poisson	$\Phi(\boldsymbol{f}, \boldsymbol{m}) = -\sum_j \frac{f_j}{m_j} \ln \frac{f_j}{m_j} + (f_j - m_j)$	Kulback
Gamma	$\Phi(\boldsymbol{f}, \boldsymbol{m}) = \sum_j \ln \frac{f_j}{m_j} + \frac{f_j}{m_j}$	Burg entropy

- **Relation with regularization:**

$$
\begin{aligned}
\widehat{\boldsymbol{f}} &= \arg\min_{\boldsymbol{f}} \left\{\Phi(\boldsymbol{f},\boldsymbol{m})\right\}, \quad \text{subject to} \quad g_i = \sum_j h_{ij} f_j \\
\widehat{\boldsymbol{f}} &= \arg\min_{\boldsymbol{f}} \left\{\Phi(\boldsymbol{f},\boldsymbol{m})\right\}, \quad \text{subject to} \quad \left\| g_i - \sum_j h_{ij}\, f_j \right\|^2 \leq \epsilon \\
\widehat{\boldsymbol{f}} &= \arg\min_{\boldsymbol{f}} \left\{\Phi(\boldsymbol{f},\boldsymbol{m}) + \alpha \|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{f}\|^2\right\} \\
&= \arg\min_{\boldsymbol{f}} \left\{\|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{f}\|^2 + \lambda \Phi(\boldsymbol{f},\boldsymbol{m})\right\}
\end{aligned}
$$

Bayesian estimation approach

- Data observation and noise models: $\longrightarrow \quad p(\boldsymbol{g}|\boldsymbol{f};\boldsymbol{\theta}_\epsilon)$
- Prior information on $\boldsymbol{f}$: $\longrightarrow \quad p(\boldsymbol{f}|\boldsymbol{\theta}_f)$
- Bayes rule:

$$p(\boldsymbol{f}|\boldsymbol{g};\boldsymbol{\theta}) = \frac{p(\boldsymbol{g}|\boldsymbol{f};\boldsymbol{\theta}_\epsilon)\, p(\boldsymbol{f}|\boldsymbol{\theta}_f)}{p(\boldsymbol{g};\boldsymbol{\theta})}$$

$$\text{where} \qquad \boldsymbol{\theta} = (\boldsymbol{\theta}_\epsilon, \boldsymbol{\theta}_f) \quad \text{and} \quad p(\boldsymbol{g}|\boldsymbol{\theta}) = \int p(\boldsymbol{g}|\boldsymbol{f};\boldsymbol{\theta}_\epsilon)\, p(\boldsymbol{f}|\boldsymbol{\theta}_f)\, \mathrm{d}\boldsymbol{f}$$

- Inference or estimation rule: cost function $c(\boldsymbol{f}, \widehat{\boldsymbol{f}})$

$$\widehat{\boldsymbol{f}} = \arg\min_{\boldsymbol{z}} \left\{ \int c(\boldsymbol{f}, \boldsymbol{z})\, p(\boldsymbol{f}|\boldsymbol{g};\boldsymbol{\theta})\, \mathrm{d}\boldsymbol{f} \right\}$$

Example:

Maximum A Posteriori (MAP):

$$\widehat{\boldsymbol{f}} = \arg\min_{\boldsymbol{f}} \{-\ln p(\boldsymbol{f}|\boldsymbol{g};\boldsymbol{\theta})\} = \arg\min_{\boldsymbol{f}} \{-\ln p(\boldsymbol{g}|\boldsymbol{f};\boldsymbol{\theta}_\epsilon) - \ln \pi_0(\boldsymbol{f}|\boldsymbol{\theta}_f)\}$$

Point Estimators:

- Maximum *a posteriori* (MAP):

$$c(\boldsymbol{f}, \widehat{\boldsymbol{f}}) = 1 - \delta(\boldsymbol{f} - \widehat{\boldsymbol{f}}) \longrightarrow \widehat{\boldsymbol{f}} = \arg\max_{\boldsymbol{f}} \{p(\boldsymbol{f}|\boldsymbol{g};\boldsymbol{\theta})\}$$

- Posterior mean (PM):

$$c(\boldsymbol{f}, \widehat{\boldsymbol{f}}) = [\boldsymbol{f} - \widehat{\boldsymbol{f}}]^t \boldsymbol{Q} [\boldsymbol{f} - \widehat{\boldsymbol{f}}]^t \longrightarrow \widehat{\boldsymbol{f}} = \mathrm{E}\{\boldsymbol{f}\} = \int \boldsymbol{f}\, p(\boldsymbol{f}|\boldsymbol{g};\boldsymbol{\theta})\,\mathrm{d}\boldsymbol{f}$$

- Marginal MAP (MMAP):

$$c(\boldsymbol{f}, \widehat{\boldsymbol{f}}) = \prod_i 1 - \delta(f_i - \widehat{f}_i) \longrightarrow \widehat{f}_i = \arg\max_{f_i} \{p_i(f_i|\boldsymbol{g};\boldsymbol{\theta}\},$$

where

$$p_i(f_i|\boldsymbol{g};\boldsymbol{\theta}) = \int p(\boldsymbol{f}|\boldsymbol{g};\boldsymbol{\theta})\,\mathrm{d}f_1 \cdots \mathrm{d}f_{i-1}\,.\,\mathrm{d}f_{i+1} \cdots \mathrm{d}f_n$$

Case of linear inverse problems

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

- Hypothesis on noise: $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, 1/\beta \boldsymbol{I}) \longrightarrow \boldsymbol{g}|\boldsymbol{f} \sim \mathcal{N}(\boldsymbol{H}\boldsymbol{f}, 1/\beta \boldsymbol{I})$

$$p(\boldsymbol{g}|\boldsymbol{f}, \beta) \propto \exp\left[-\frac{\beta}{2}\|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{f}\|^2\right]$$

- Hypothesis on $\boldsymbol{f}$: $\quad \boldsymbol{f} \sim \mathcal{N}(\boldsymbol{f}_0, 1/\alpha \boldsymbol{P}_0)$

$$p(\boldsymbol{f}|\alpha) \propto \exp\left[-\frac{\alpha}{2}[\boldsymbol{f} - \boldsymbol{f}_0]^t \boldsymbol{P}_0^{-1}[\boldsymbol{f} - \boldsymbol{f}_0]\right]$$

- Posterior $\quad p(\boldsymbol{f}|\boldsymbol{g}) \sim \mathcal{N}(\widehat{\boldsymbol{f}}, \widehat{\boldsymbol{P}})$ with

$$\widehat{\boldsymbol{f}} = \widehat{\boldsymbol{P}}\boldsymbol{H}^t(\boldsymbol{g} - \boldsymbol{H}\boldsymbol{f}_0), \quad \widehat{\boldsymbol{P}} = \left(\boldsymbol{H}^t\boldsymbol{H} + \lambda \boldsymbol{P}_0^{-1}\right)^{-1}, \quad \lambda = \frac{\alpha}{\beta}$$

- MAP estimate:

$$\widehat{\boldsymbol{f}} = \arg\min_{\boldsymbol{f}} \{J(\boldsymbol{f}) = Q(\boldsymbol{f}) + \lambda\Phi(\boldsymbol{f})\} \text{ with}$$

$$Q(\boldsymbol{f}) = \|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{f}\|^2, \quad \Phi(\boldsymbol{f}) = \boldsymbol{f}^t\boldsymbol{P}_0^{-1}\boldsymbol{f} = \|\boldsymbol{D}\boldsymbol{f}\|^2, \quad \lambda = \frac{\alpha}{\beta}$$

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

- Gamma law hypothesis for $\boldsymbol{f}$:

$$p(f_j) \propto (f_j/m_j)^\alpha \exp\left[-f_j/m_j\right] \longrightarrow \quad \Phi_\alpha(\boldsymbol{f}) = \alpha \sum_j \ln \frac{f_j}{m_j} + \frac{f_j}{m_j}$$

- Beta law hypothesis for $\boldsymbol{f}$:

$$p(f_j) \propto f_j^\alpha (1-f_j)^{1-\alpha} \longrightarrow \quad \Phi_\alpha(\boldsymbol{f}) = \alpha \sum_j \ln f_j + (1-\alpha) \sum_j \ln(1-f_j)$$

- Generalized Gaussian laws for $\boldsymbol{f}$:

$$p(f_j) \propto \exp\left[-\alpha|f_j|^p\right], \quad 1 < p < 2 \longrightarrow \quad \Phi_\alpha(\boldsymbol{f}) = \alpha \sum_j |f_j|^p$$

- Markovian models for $\boldsymbol{f}$:

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

General MAP estimate:

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

- Gaussian laws: $\Phi_\alpha(\boldsymbol{f})$ quadratic
  $\longrightarrow J(\boldsymbol{f})$ quadratic $\longrightarrow \widehat{\boldsymbol{f}}$ linear function of $\boldsymbol{g}$
- Entropic laws: $\Phi_\alpha(\boldsymbol{f})$ convex and separable: $\Phi(\boldsymbol{f}) = \sum_j \phi_\alpha(f_j)$

with $\quad \phi_\alpha(f_j) = \left\{\alpha f_j^p, \quad \alpha(f_j \ln f_j - f_j), \quad \alpha(\ln f_j - f_j), \ldots\right\}$
  $\longrightarrow J(\boldsymbol{f})$ convex $\longrightarrow$ Nonlinear estimators but easy to calculate

- Markovian models: $\Phi_\alpha(\boldsymbol{f}) = \sum_j \sum_{i\in N_j} \phi_\alpha(f_j - f_i)$ with $\phi_\alpha(t) =$

$$\left\{\begin{array}{ll} |t|^2 & \text{if } |t| < \alpha, \\ \alpha^2 & \text{else,} \end{array}\right. , \left\{\begin{array}{ll} t^2 & \text{if } |t| < \alpha, \\ 2\alpha t - \alpha^2 & \text{else,} \end{array}\right. , \frac{\alpha^2 t^2}{1+t^2}, \quad \log\cosh(t/\alpha)$$

$\Phi_\alpha(\boldsymbol{f})$ non convex $\longrightarrow$ Nonlinear estimator and
need of global optimization

Open problems

- Choice of $p(\boldsymbol{f}|\boldsymbol{\theta}_f)$: Gaussian / Entropic / Markovian
  Markovian models: choice of the potential functions
  $\longrightarrow$ Transformation group invariance, ME,
  Scale invariance, etc. (Djafari93, Brette *et al.* 93, 94)

- Hyperparameters estimation and model selection
  $\longrightarrow$ GMAP, MAPM, EM, SEM, ...

- Choice of the estimation criteria : MAP, PM or MMAP
  – MAP : Optimization
  – PM or MMAP : Integration

- Optimization and integration algorithms and implementation
  – Local / Global, Determinist / Stochastic
  – multigrids & multiresolution

Multi sensor image processing problems

- Disjoint multi sensors system:

$$\boldsymbol{g}_i = \boldsymbol{H}_i \boldsymbol{f}_i + \boldsymbol{\epsilon}_i, \quad i = 1, \cdots, M$$

- General multi input multi output (MIMO) system:

$$\boldsymbol{g}_i = \sum_{j=1}^{N} \boldsymbol{H}_{ij} \boldsymbol{f}_j + \boldsymbol{\epsilon}_i, \quad i = 1, \cdots, M$$

- General unknown mixing gain MIMO system:

$$\boldsymbol{g}_i = \sum_{j=1}^{N} A_{ij} \boldsymbol{H}_j \boldsymbol{f}_j + \boldsymbol{\epsilon}_i, \quad i = 1, \cdots, M$$

- Blind Sources Separation (BSS) problem:

$$\boldsymbol{g}_i = \sum_{j=1}^{N} A_{ij} \boldsymbol{f}_j + \boldsymbol{\epsilon}_i, \quad i = 1, \cdots, M$$

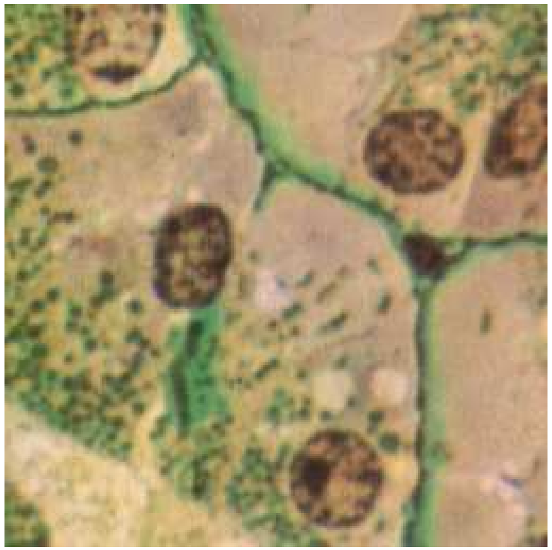
where $\boldsymbol{A} = \{A_{ij}, i = 1, \cdots, M, \ j = 1, \cdots, N\}$ is an unknown mixing matrix.

Multi-spectral image deconvolution

$$\epsilon_i(x,y)$$

$$f_i(x,y) \longrightarrow \boxed{h(x,y)} \longrightarrow \oplus \longrightarrow g_i(x,y) = f_i(x,y) * h(x,y) + \epsilon_i(x,y)$$

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



?

$\Longleftarrow$

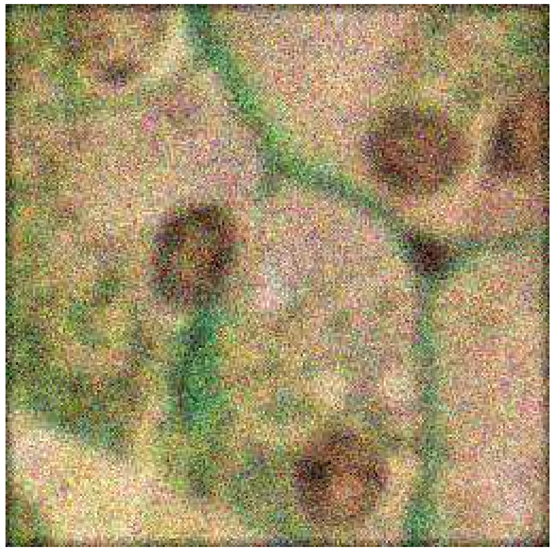
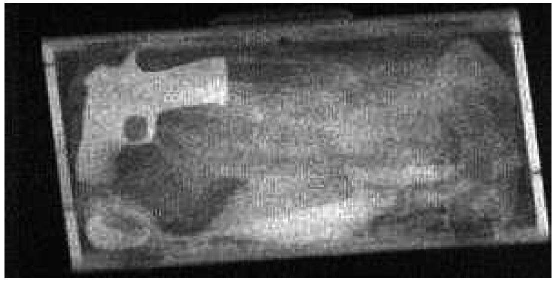
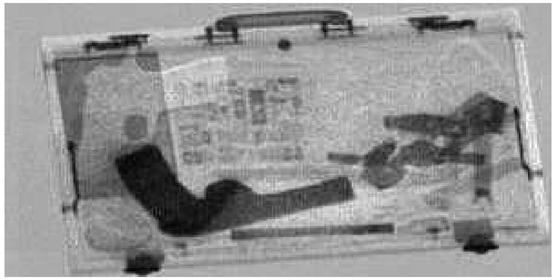


Image fusion and joint segmentation



Fusion ?

$$\Longrightarrow$$

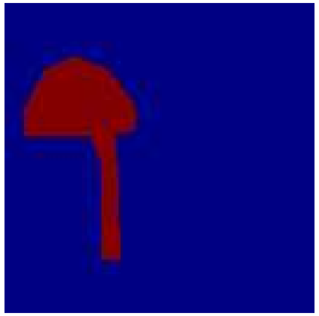


$$g_i(\boldsymbol{r}) = f_i(\boldsymbol{r}) + \epsilon_i(\boldsymbol{r}), \quad i = 1, \cdots, M$$

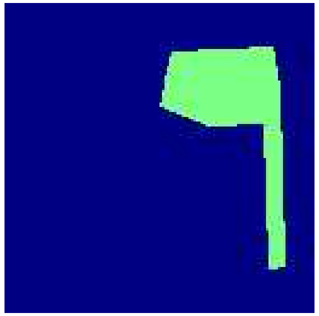
$$\underline{g}(\boldsymbol{r}) = \{g_i(\boldsymbol{r}),\ i = 1, M\}, \qquad \boldsymbol{g}_i = \{g_i(\boldsymbol{r}),\ \boldsymbol{r} \in \mathcal{R}\}, \qquad \underline{\boldsymbol{g}} = \{\boldsymbol{g}_i(\boldsymbol{r}),\ i = 1, M\}$$

$$\underline{g}(\boldsymbol{r}) = \underline{f}(\boldsymbol{r}) + \underline{\epsilon}(\boldsymbol{r}), \qquad\qquad \underline{\boldsymbol{g}} = \underline{\boldsymbol{f}} + \underline{\boldsymbol{\epsilon}}$$

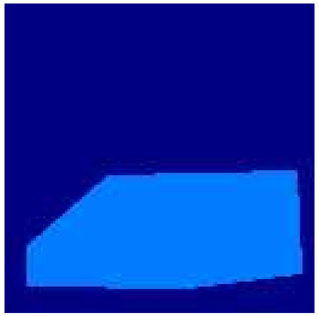
Blind image separation and joint segmentation



$f_1(\boldsymbol{r})$



$f_2(\boldsymbol{r})$

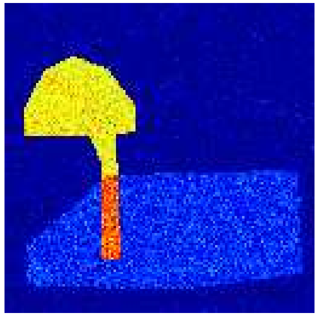


$f_3(\boldsymbol{r})$

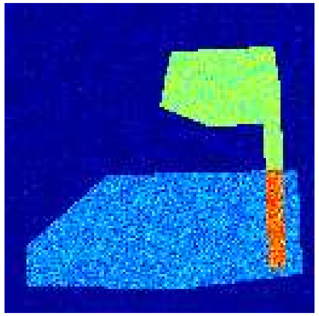
?

Separation

$\Longleftarrow$



$g_1(\boldsymbol{r})$



$g_2(\boldsymbol{r})$

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

$$\underline{g}(\boldsymbol{r}) = \{g_i(\boldsymbol{r}),\; i = 1, M\}$$

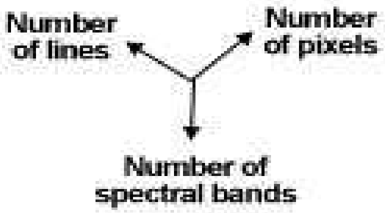
$$\underline{g}(\boldsymbol{r}) = \boldsymbol{A}\underline{f}(\boldsymbol{r}) + \underline{\epsilon}(\boldsymbol{r}),$$

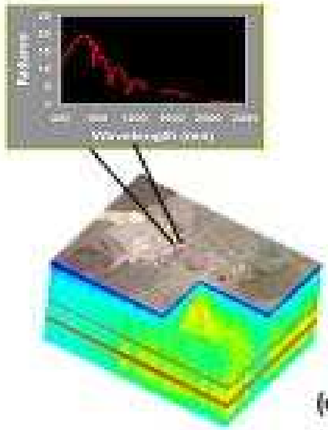
$$\underline{\boldsymbol{g}} = \{\boldsymbol{g}_i(\boldsymbol{r}),\; i = 1, M\}$$

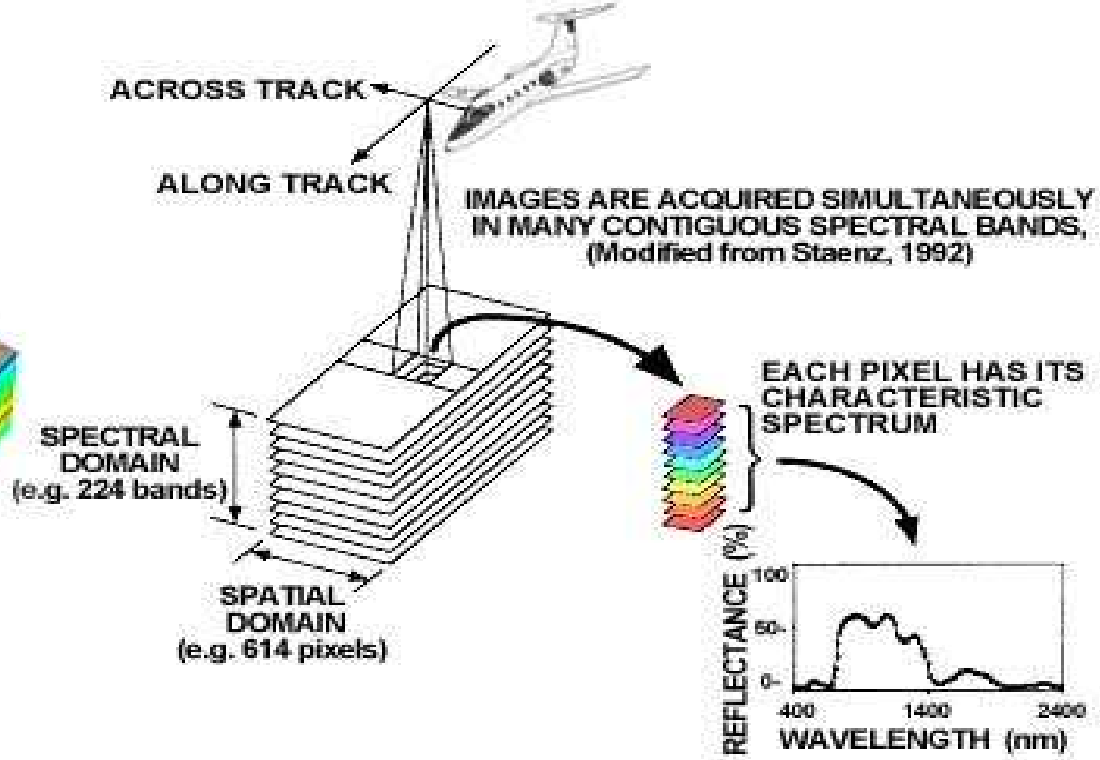
$$\boldsymbol{g}_i = \{g_i(\boldsymbol{r}),\; \boldsymbol{r} \in \mathcal{R}\},$$

$$\underline{\boldsymbol{g}} = \boldsymbol{A}\underline{\boldsymbol{f}} + \underline{\boldsymbol{\epsilon}}$$

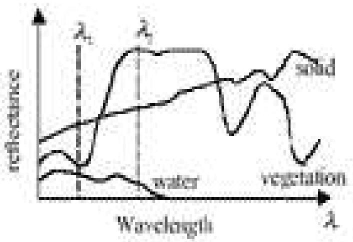
Segmentation of hyperspectral images



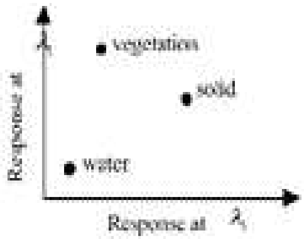




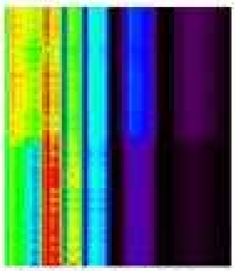
(a) Image Space



(b) Spectral Space

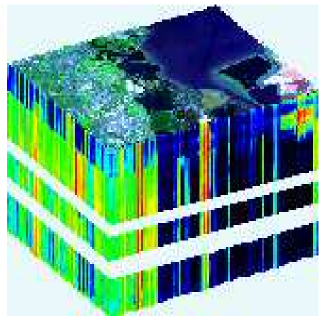


(c) Feature Space

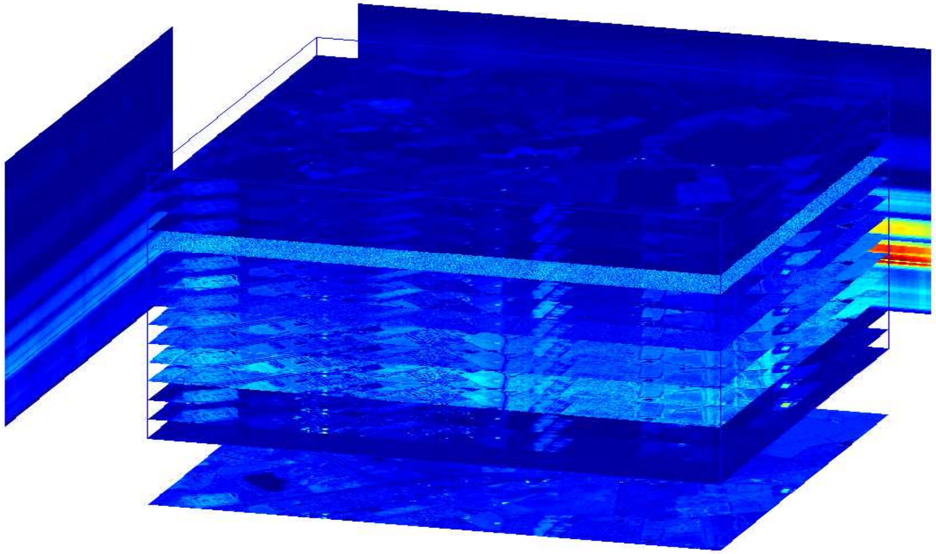


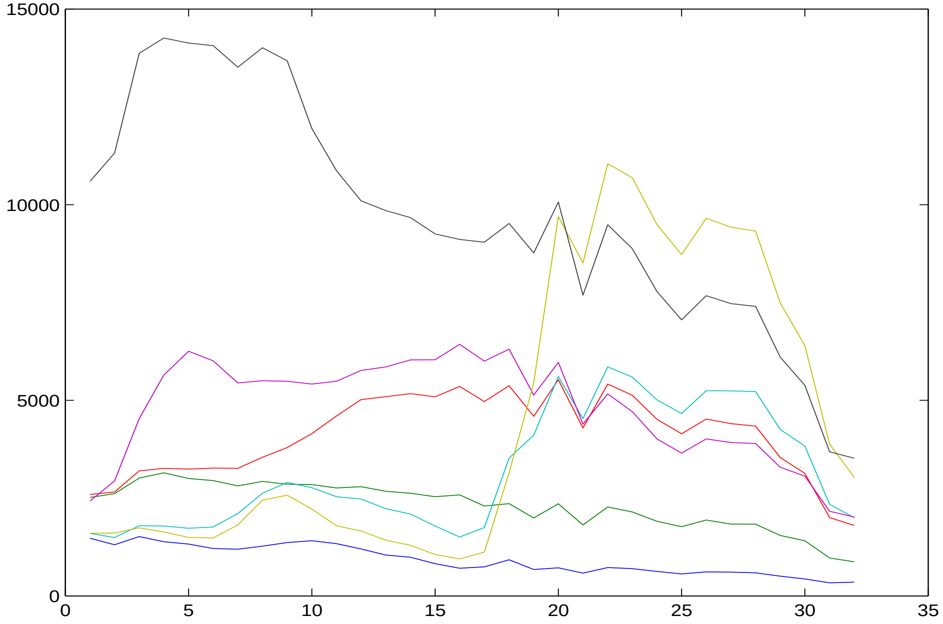
(a) Spectral slice

(b) 3D image cube



Segmentation of hyperspectral images



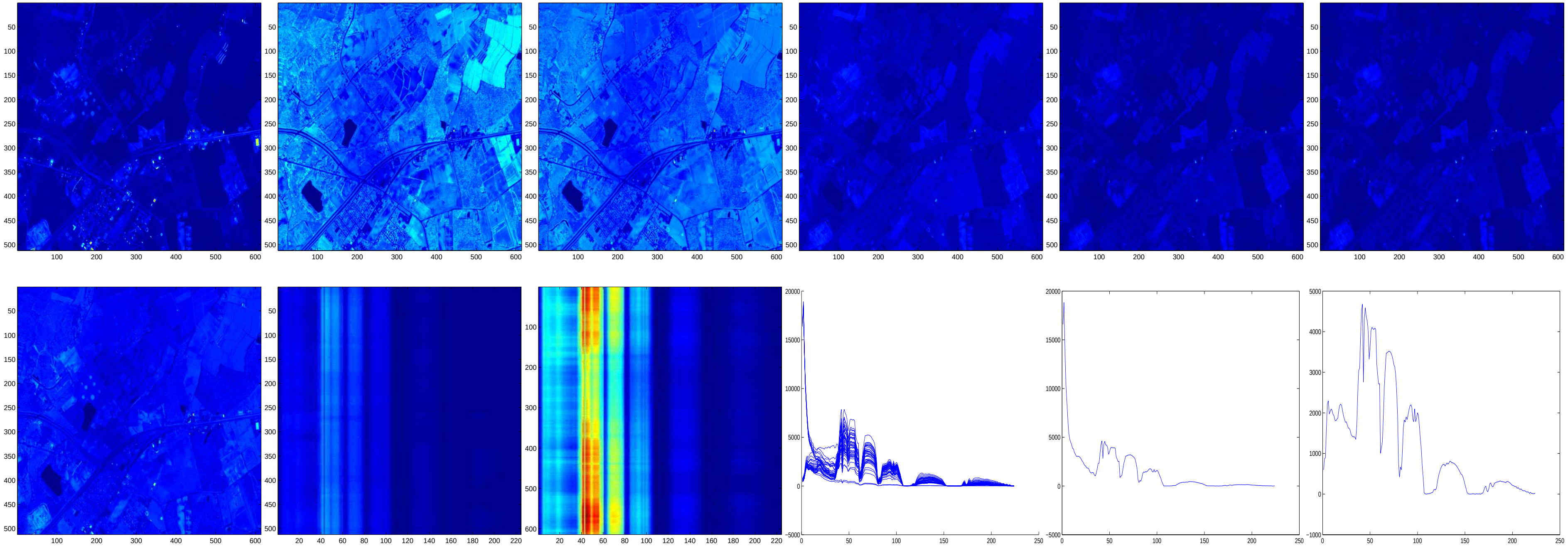


$$g_i(\boldsymbol{r}) = f_i(\boldsymbol{r}) + \epsilon_i(\boldsymbol{r}), \quad i = 1, \cdots, M$$

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

$$\underline{g}(\boldsymbol{r}) = \underline{f}(\boldsymbol{r}) + \underline{\epsilon}(\boldsymbol{r}), \qquad \underline{\boldsymbol{g}} = \underline{\boldsymbol{f}} + \underline{\boldsymbol{\epsilon}}$$

Segmentation of hyperspectral images

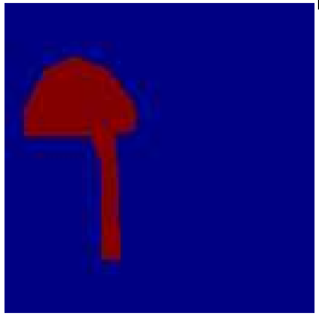


$$g_i(\boldsymbol{r}) = f_i(\boldsymbol{r}) + \epsilon_i(\boldsymbol{r}), \quad i = 1, \cdots, M$$

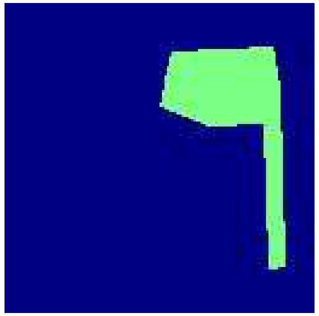
$$\underline{g}(\boldsymbol{r}) = \{g_i(\boldsymbol{r}),\ i = 1, M\}, \qquad \boldsymbol{g}_i = \{g_i(\boldsymbol{r}),\ \boldsymbol{r} \in \mathcal{R}\}, \qquad \underline{\boldsymbol{g}} = \{\boldsymbol{g}_i(\boldsymbol{r}),\ i = 1, M\}$$

$$\underline{g}(\boldsymbol{r}) = \underline{f}(\boldsymbol{r}) + \underline{\epsilon}(\boldsymbol{r}), \qquad \underline{\boldsymbol{g}} = \underline{\boldsymbol{f}} + \underline{\boldsymbol{\epsilon}}$$

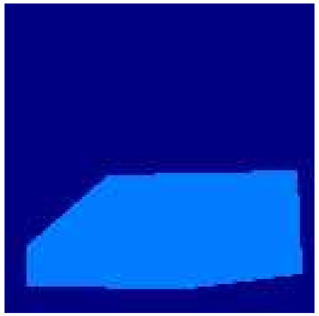
Blind image separation and joint segmentation



$f_1(\boldsymbol{r})$



$f_2(\boldsymbol{r})$

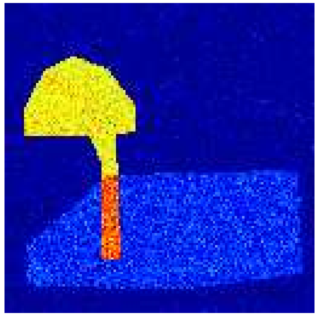


$f_3(\boldsymbol{r})$

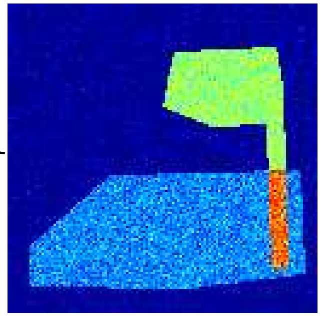
?

Separation

$\Longleftarrow$



$g_1(\boldsymbol{r})$



$g_2(\boldsymbol{r})$

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

$$\underline{g}(\boldsymbol{r}) = \{g_i(\boldsymbol{r}),\ i = 1, M\}$$

$$\underline{g}(\boldsymbol{r}) = \boldsymbol{A}\underline{f}(\boldsymbol{r}) + \underline{\epsilon}(\boldsymbol{r}),$$

$$\underline{\boldsymbol{g}} = \{\boldsymbol{g}_i(\boldsymbol{r}),\ i = 1, M\}$$

$$\boldsymbol{g}_i = \{g_i(\boldsymbol{r}),\ \boldsymbol{r} \in \mathcal{R}\},$$

$$\underline{\boldsymbol{g}} = \boldsymbol{A}\underline{\boldsymbol{f}} + \underline{\boldsymbol{\epsilon}}$$

Bayesian estimation approach

$$\underline{\boldsymbol{g}} = \boldsymbol{A}\underline{\boldsymbol{f}} + \underline{\boldsymbol{\epsilon}}$$

- Forward model and prior knowledge on the noise $\longrightarrow$ $p(\underline{\boldsymbol{g}}|\underline{\boldsymbol{f}})$
- Prior knowledge on $\underline{\boldsymbol{f}}$ $\longrightarrow$ $p(\underline{\boldsymbol{f}})$
- Bayes rule: $p(\underline{\boldsymbol{f}}|\underline{\boldsymbol{g}}) = p(\underline{\boldsymbol{g}}|\underline{\boldsymbol{f}})\; p(\underline{\boldsymbol{f}}) \,/\, p(\underline{\boldsymbol{g}}) \propto p(\underline{\boldsymbol{g}}|\underline{\boldsymbol{f}})\; p(\underline{\boldsymbol{f}})$
- Infer on $\underline{\boldsymbol{f}}$ via $p(\underline{\boldsymbol{f}}|\underline{\boldsymbol{g}})$:
  - MAP estimator:

$$\widehat{\underline{\boldsymbol{f}}} = \arg\max_{\underline{\boldsymbol{f}}} \left\{p(\underline{\boldsymbol{f}}|\underline{\boldsymbol{g}})\right\} = \arg\min_{\boldsymbol{f}} \left\{-\ln p(\underline{\boldsymbol{f}}|\underline{\boldsymbol{g}})\right\}$$

  - MSE estimator $\quad \widehat{\underline{\boldsymbol{f}}} = \int \underline{\boldsymbol{f}}\; p(\underline{\boldsymbol{f}}|\underline{\boldsymbol{g}}) \,\mathrm{d}\underline{\boldsymbol{f}}$ posterior mean

Bayesian blind sources separation (BBSS)

$$\underline{\boldsymbol{g}} = \boldsymbol{A}\underline{\boldsymbol{f}} + \underline{\boldsymbol{\epsilon}}$$

- Likelihood $\quad p(\underline{\boldsymbol{g}}|\underline{\boldsymbol{f}}, \boldsymbol{A}, \underline{\boldsymbol{\theta}}_1)$
- Priors $\quad p(\underline{\boldsymbol{f}}|\underline{\boldsymbol{\theta}}_2)$, $p(\boldsymbol{A}|\boldsymbol{\theta}_3)$ and $p(\boldsymbol{\theta})$ $\quad$ hyperparameters $\boldsymbol{\theta} = \{\underline{\boldsymbol{\theta}}_1\, \underline{\boldsymbol{\theta}}_2, \boldsymbol{\theta}_3\}$
- Bayes rule

$$p(\underline{\boldsymbol{f}}, \boldsymbol{A}, \boldsymbol{\theta}|\underline{\boldsymbol{g}}) \propto p(\underline{\boldsymbol{g}}|\underline{\boldsymbol{f}}, \boldsymbol{A}, \underline{\boldsymbol{\theta}}_1)\, p(\underline{\boldsymbol{f}}|\underline{\boldsymbol{\theta}}_2)\, p(\boldsymbol{A}|\boldsymbol{\theta}_3)\, p(\boldsymbol{\theta})$$

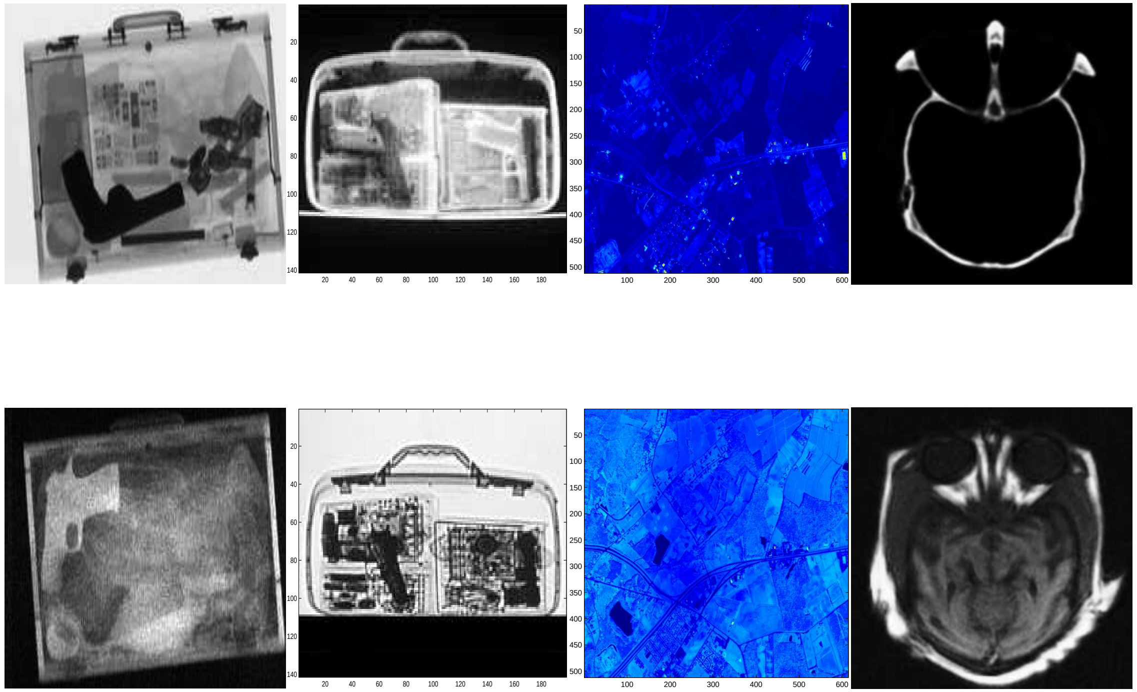
- Generalized a posteriori EM-ML estimator

$$(\widehat{\boldsymbol{A}}, \widehat{\boldsymbol{\theta}}) = \arg\max_{(\boldsymbol{A},\boldsymbol{\theta})} \{p(\boldsymbol{A}, \boldsymbol{\theta}|\underline{\boldsymbol{g}})\} \quad \longrightarrow \quad \widehat{\underline{\boldsymbol{f}}} = \arg\max_{\underline{\boldsymbol{f}}} \left\{p(\underline{\boldsymbol{f}}|\underline{\boldsymbol{g}}, \widehat{\boldsymbol{A}}, \widehat{\boldsymbol{\theta}})\right\}$$

- MSE estimator which corresponds to the posterior mean using MCMC

$$\{\widehat{\underline{\boldsymbol{f}}}, \widehat{\boldsymbol{A}}, \widehat{\boldsymbol{\theta}}\} = \int \{\widehat{\underline{\boldsymbol{f}}}, \widehat{\boldsymbol{A}}, \widehat{\boldsymbol{\theta}}\}\, p(\underline{\boldsymbol{f}}, \boldsymbol{A}, \boldsymbol{\theta}|\underline{\boldsymbol{g}})\, \mathrm{d}\{\widehat{\underline{\boldsymbol{f}}}, \widehat{\boldsymbol{A}}, \widehat{\boldsymbol{\theta}}\}.$$

HMM modeling of images



What they share ?

Borders & Regions

Segmentation

Hidden variable $z(\boldsymbol{r})$

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

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

Homogeneity in regions

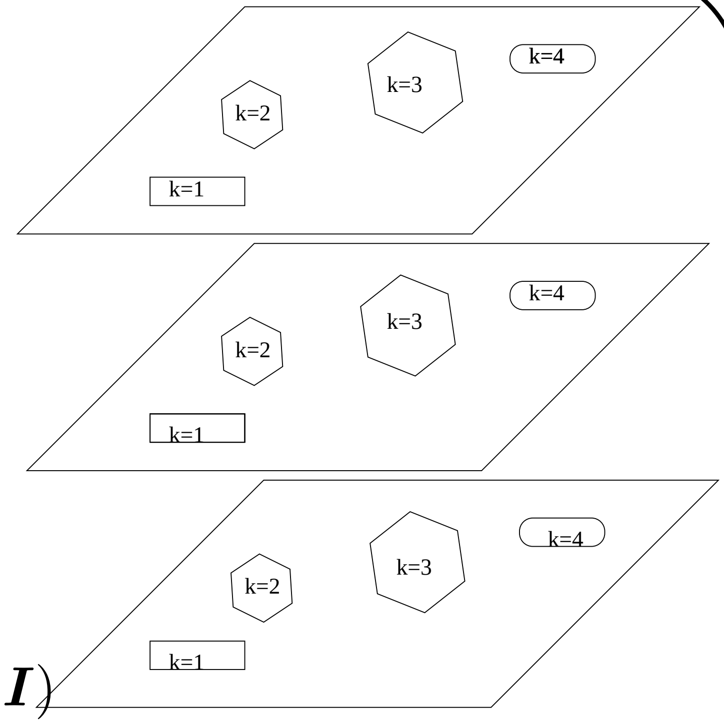
$p(f(\boldsymbol{r})|z(\boldsymbol{r}) = k) = \mathcal{N}(m_k, \sigma_k^2)$

$$R_{j\,k} = \{\boldsymbol{r} \;:\; z_j(\boldsymbol{r}) = k\}, \qquad \mathcal{R}_j = \cup_k R_{j\,k}$$

$$\boldsymbol{f}_{j\,k} = \{f_j(\boldsymbol{r}) \;:\; \boldsymbol{r} \in R_{j\,k}\}, \quad \boldsymbol{f}_j = \cup_k \boldsymbol{f}_{j\,k}.$$

$$p(f_j(\boldsymbol{r}), \boldsymbol{r} \in \mathcal{R}) = \prod_{k=1}^{K_j} p(f_j(\boldsymbol{r}), \boldsymbol{r} \in R_{j\,k}) = \prod_{k=1}^{K_j} \mathcal{N}(m_{j\,k}\mathbf{1}, \sigma_{j\,k}^2)$$

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.: $p(f_j(\boldsymbol{r})|z_j(\boldsymbol{r}) = k) = \mathcal{N}(m_{j\,k}, \sigma^2_{j\,k})$
    $$p(f_j(\boldsymbol{r}), \boldsymbol{r} \in R_{j\,k}) = \mathcal{N}(m_{j\,k}\mathbf{1}, \sigma^2_{j\,k}\boldsymbol{I})$$
  - Markovien: $p(f_j(\boldsymbol{r}), \boldsymbol{r} \in R_{j\,k}) = \mathcal{N}(m_{j\,k}\mathbf{1}, \boldsymbol{\Sigma}_{j\,k})$
- For pixels values in different images but in a given common region two possibilities:
  - i.i.d.: $(f_j(\boldsymbol{r})|z(\boldsymbol{r}) = k)$ independent of $(f_i(\boldsymbol{r})|z(\boldsymbol{r}) = k), \quad i \neq j$
  - Markovien: $p(f_j(\boldsymbol{r})|z(\boldsymbol{r}) = k, \underline{\boldsymbol{f}}(\boldsymbol{r})) = p(f_j(\boldsymbol{r})|z(\boldsymbol{r}) = k, f_{j-1}(\boldsymbol{r}))$

Modeling the labels

$$p(f_j(\boldsymbol{r})|z_j(\boldsymbol{r})=k)=\mathcal{N}(m_{j\,k},\sigma_{j\,k}^2)\longrightarrow p(f_j(\boldsymbol{r}))=\sum_k P(z_j(\boldsymbol{r})=k)\,\mathcal{N}(m_{j\,k},\sigma_{j\,k}^2)$$

- Independent Gaussian Mixture model (IGM): $\boldsymbol{z}_j=\{z_j(\boldsymbol{r}),\boldsymbol{r}\in\mathcal{R}\}$ i.i.d.

$$P(z_j(\boldsymbol{r})=k)=p_k,\quad \text{with} \quad \sum_k p_k=1 \;\text{ and }\; p(\boldsymbol{z}_j)=\prod_k p_k$$

- Contextual Gaussian Mixture model (CGM): $\boldsymbol{z}_j$ Markovien

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

which is the Potts Markov random feild (PMRF).

The parameter α controls the mean value of the regions' sizes.

Expressions of likelihood, prior and posterior laws

$$\boldsymbol{g}_i = \sum_{j=1}^{N} A_{i,j} \boldsymbol{f}_j + \boldsymbol{\epsilon}_i, \quad i = 1, \cdots, M \quad \longrightarrow \quad \underline{\boldsymbol{g}} = \boldsymbol{A} \underline{\boldsymbol{f}} + \underline{\boldsymbol{\epsilon}}$$

- Likelihood:

$$p(\underline{\boldsymbol{g}} | \boldsymbol{A}, \underline{\boldsymbol{f}}, \boldsymbol{\theta}_1) = \prod_{i=1}^{M} p(\boldsymbol{g}_i | \boldsymbol{A}, \underline{\boldsymbol{f}}, \boldsymbol{\Sigma}_{\epsilon i}) = \prod_{i=1}^{M} \mathcal{N}(\boldsymbol{A} \underline{\boldsymbol{f}}, \boldsymbol{\Sigma}_{\epsilon i})$$

$\boldsymbol{\Sigma}_{\epsilon i} = \sigma_{\epsilon}{}_i^2 \boldsymbol{I}, \quad \longrightarrow \quad \boldsymbol{\theta}_1 = \{\sigma_{\epsilon}{}_i^2, i = 1, \cdots, M\}$

- HMM for the images:

$$p(\underline{\boldsymbol{f}} | \underline{\boldsymbol{z}}, \boldsymbol{\theta}_2) = \prod_{j=1}^{N} p(\boldsymbol{f}_j | \boldsymbol{z}_j, \boldsymbol{m}_j, \boldsymbol{\Sigma}_j)$$

where $\underline{\boldsymbol{z}} = \{\boldsymbol{z}_j, j = 1, \cdots, N\}$ and where we assumed that $\boldsymbol{f}_j | \boldsymbol{z}_j$ are independent. $\longrightarrow \boldsymbol{\theta}_2 = \{(m_{ik}, \sigma_{j\,k}^2), j = 1, \cdots, N\}$

- PMRF for the labels:

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

- Conjugate priors for the hyperparameters $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$:
  $\boldsymbol{\theta} = \{\{\sigma_{\epsilon_i}^2, i = 1, \cdots, M\}, \{(m_{i\,k}, \sigma_{j\,k}^2), j = 1, \cdots, M, k = 1, \cdots, K\}\}$

$$\begin{aligned}
p(m_{j\,k}) &= \mathcal{N}(m_{j\,k0}, \sigma_{j\,k0}^2) \\
p(\sigma_{j\,k}^2) &= \mathcal{IG}(\alpha_{j0}, \beta_{j0}) \\
p(\boldsymbol{\Sigma}_{j\,k}) &= \mathcal{IW}(\alpha_{j0}, \Lambda_{j0}) \\
p(\sigma_{\epsilon\,i}) &= \mathcal{IG}(\alpha_{i0}, \beta_{i0})
\end{aligned}$$

- Joint posterior law of $\underline{\boldsymbol{f}}$, $\underline{\boldsymbol{z}}$ and $\underline{\boldsymbol{\theta}}$

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

General MCMC sampling scheme

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

Gibbs sampling:

- Generate samples $(\underline{\boldsymbol{f}}, \underline{\boldsymbol{z}}, \underline{\boldsymbol{\theta}})^{(1)}, \cdots, (\underline{\boldsymbol{f}}, \underline{\boldsymbol{z}}, \underline{\boldsymbol{\theta}})^{(N)}$ using
  - $\underline{\boldsymbol{f}} \sim p(\underline{\boldsymbol{f}} | \underline{\boldsymbol{g}}, \underline{\boldsymbol{z}}, \underline{\boldsymbol{\theta}}),$
  - $\underline{\boldsymbol{z}} \sim p(\underline{\boldsymbol{z}} | \underline{\boldsymbol{g}}, \underline{\boldsymbol{f}}, \underline{\boldsymbol{\theta}}),$
  - $\underline{\boldsymbol{\theta}} \sim p(\underline{\boldsymbol{\theta}} | \underline{\boldsymbol{g}}, \underline{\boldsymbol{f}}, \underline{\boldsymbol{z}}),$
- Compute any statistics such as mean, median, variance, ...

Difficulties:

- No mixture, No convolution: $\boldsymbol{g}_i = \boldsymbol{f}_i + \boldsymbol{\epsilon}_i, \quad i = 1, \cdots, M$
- Mixture but No convolution: $\boldsymbol{g}_i = \sum_{j=1}^{N} A_{ij} \, \boldsymbol{f}_j + \boldsymbol{\epsilon}_i, \quad i = 1, \cdots, M$
- Convolution but No mixture: $\boldsymbol{g}_i = \boldsymbol{H}_i \, \boldsymbol{f}_i + \boldsymbol{\epsilon}_i, \quad i = 1, \cdots, M$
- Mixture and Convolution: $\boldsymbol{g}_i = \sum_{j=1}^{N} A_{ij} \, \boldsymbol{H}_{ij} \boldsymbol{f}_j + \boldsymbol{\epsilon}_i, \quad i = 1, \cdots, M$

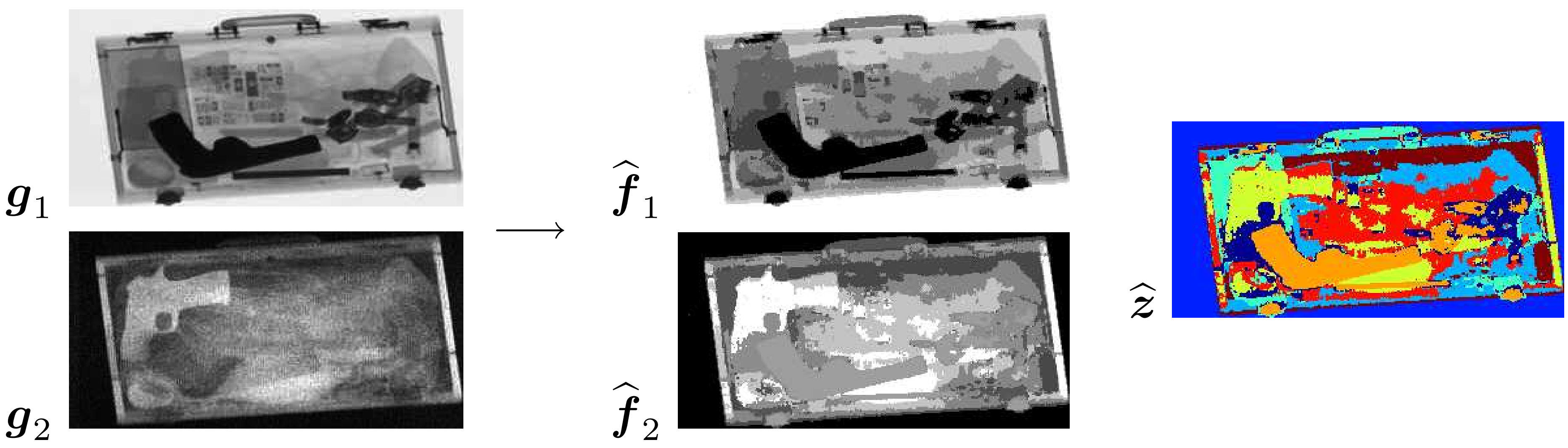
Examples of applications

- No mixture, No convolution applications:
  - Multi channel image fusion and joint segmentation
    $$\boldsymbol{g}_i = \boldsymbol{f}_i + \boldsymbol{\epsilon}_i, \quad i = 1, \cdots, M, \qquad \boldsymbol{f}_i | \boldsymbol{z} \ \text{ independent}$$
  - Hyperspectral image segmentation
    $$\boldsymbol{g}_i = \boldsymbol{f}_i + \boldsymbol{\epsilon}_i, \quad i = 1, \cdots, M, \qquad \boldsymbol{f}_i | \boldsymbol{z} \ \text{ dependent}$$
  - Video movie segmentation with motion estimation
    $$\boldsymbol{g}_i = \boldsymbol{f}_i + \boldsymbol{\epsilon}_i, \quad i = 1, \cdots, M, \qquad \boldsymbol{f}_i | \boldsymbol{z}_i \ \text{ independent}$$
- Mixture, No convolution applications:
  - Blind source (image) separation (BSS) and joint segmentation
    $\boldsymbol{g}_i = \sum_{j=1}^{N} A_{ij}\, \boldsymbol{f}_j + \boldsymbol{\epsilon}_i, \quad i = 1, \cdots, M, \quad \boldsymbol{f}_i | \boldsymbol{z}_i \ \text{ independent}$
- Convolution but No mixture applications:
  - Fourier synthesis in optical imaging $\qquad \boldsymbol{g} = \boldsymbol{H}(\boldsymbol{f}) + \boldsymbol{\epsilon}, \quad \boldsymbol{f} | \boldsymbol{z}$
  - Single channel image restoration $\qquad \boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\epsilon}, \quad \boldsymbol{f} | \boldsymbol{z}$

Images fusion and joint segmentation

(Olivier FÉRON)

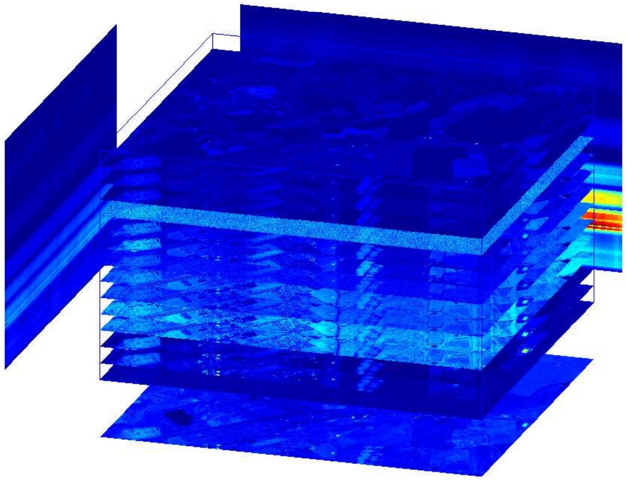
$$\begin{cases} g_i(\boldsymbol{r}) = f_i(\boldsymbol{r}) + \epsilon_i(\boldsymbol{r}) \\ p(f_i(\boldsymbol{r})|z(\boldsymbol{r}) = k) = \mathcal{N}(m_{ik}, \sigma^2_{i\,k}) \\ p(\underline{\boldsymbol{f}}|\boldsymbol{z}) = \prod_i p(\boldsymbol{f}_i|\boldsymbol{z}) \end{cases}$$

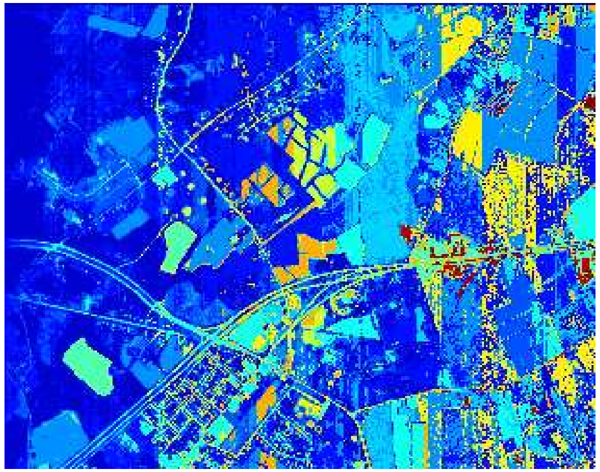


Joint segmentation of hyper-spectral images

(Adel MOHAMMADPOUR)

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

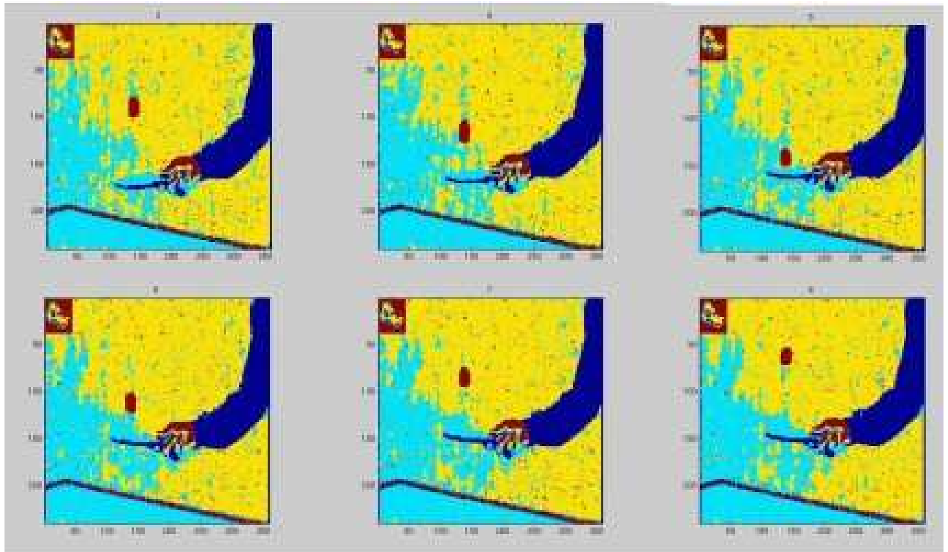




Segmentation of a video sequence of images

(Patrice BRAULT)

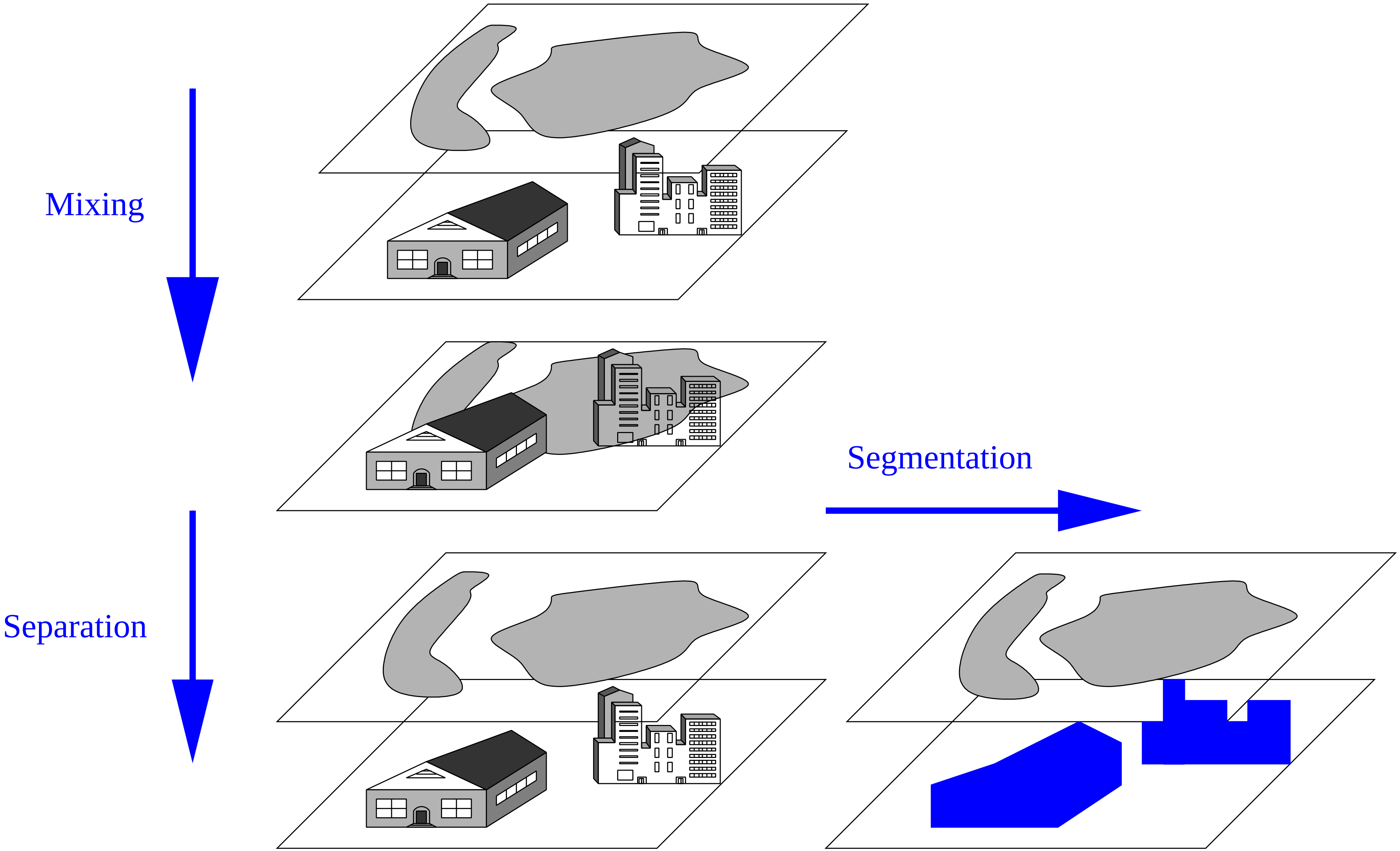
$$
\left\{
\begin{array}{l}
g_i(\boldsymbol{r}) = f_i(\boldsymbol{r}) + \epsilon_i(\boldsymbol{r}), \quad \boldsymbol{r} \in \mathcal{R}, \quad \text{or} \quad \boldsymbol{g}_i = \boldsymbol{f}_i + \boldsymbol{\epsilon}_i, \quad \text{or} \quad \underline{\boldsymbol{g}} = \underline{\boldsymbol{f}}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

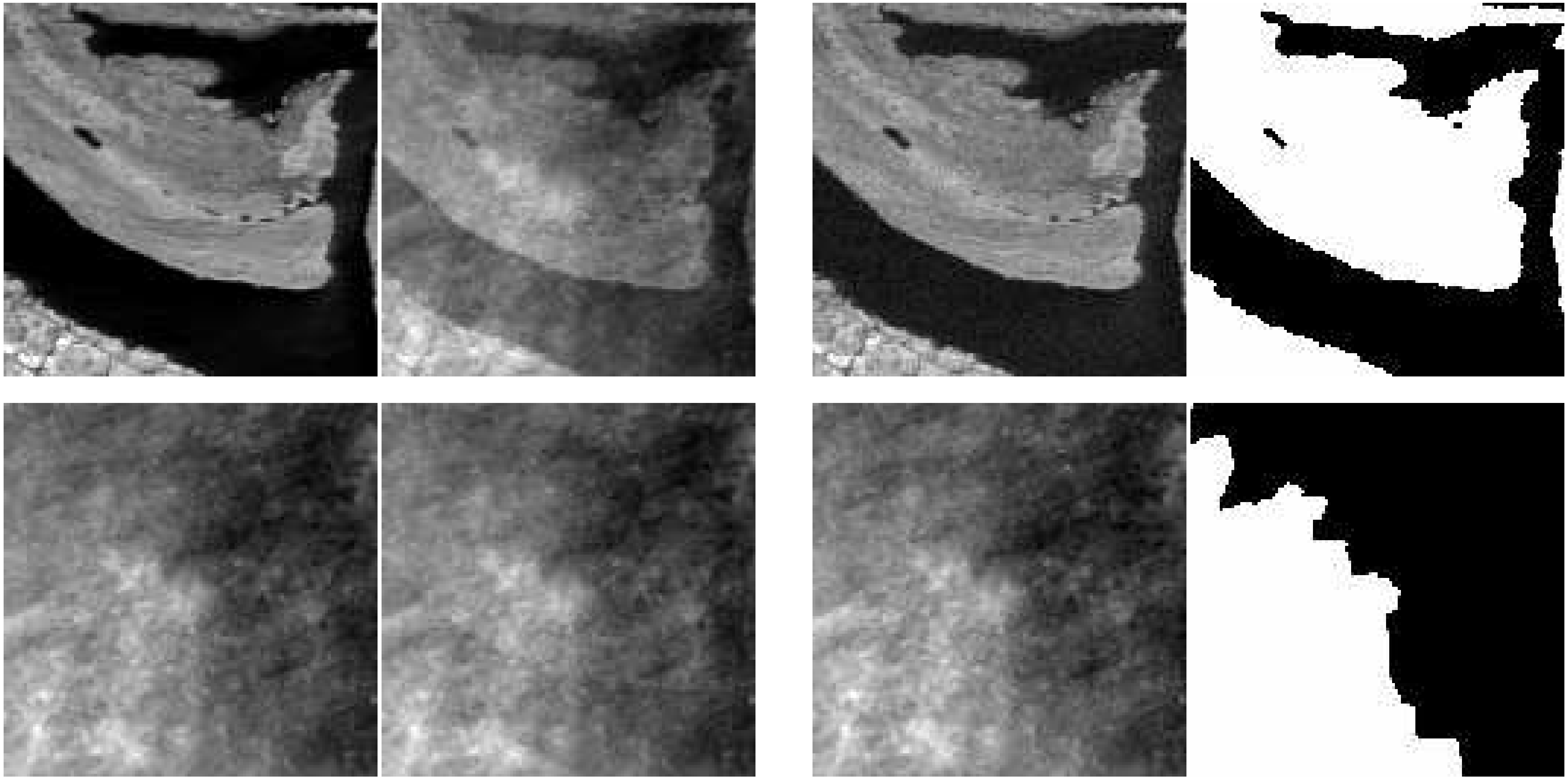


Blind image separation and joint segmentation

$$\left\{\begin{array}{l}
g_i(\boldsymbol{r}) = \sum_{j=1}^{N} A_{ij} f_j(\boldsymbol{r}) + \epsilon_i(\boldsymbol{r}) \longrightarrow \boldsymbol{g}(\boldsymbol{r}) = \boldsymbol{A}\boldsymbol{f}(\boldsymbol{r}) + \boldsymbol{\epsilon}(\boldsymbol{r}) \longrightarrow \underline{\boldsymbol{g}} = \boldsymbol{A}\underline{\boldsymbol{f}} + \underline{\boldsymbol{\epsilon}} \\
p(\underline{\boldsymbol{g}}|\underline{\boldsymbol{f}}, \boldsymbol{A}, \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}) = \prod_{\boldsymbol{r}\in\mathcal{R}} p(\boldsymbol{g}(\boldsymbol{r})|\boldsymbol{f}(\boldsymbol{r}), \boldsymbol{A}) = \prod_{\boldsymbol{r}\in\mathcal{R}} \mathcal{N}(\boldsymbol{A}\boldsymbol{f}(\boldsymbol{r}), \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}) \\
p(f_j(\boldsymbol{r})|z_j(\boldsymbol{r}) = k) = \mathcal{N}(m_{j\,k}, \sigma_{j\,k}^2), \\
p(\underline{\boldsymbol{f}}|\underline{\boldsymbol{z}}) = \prod_{\boldsymbol{r}\in\mathcal{R}} \prod_j p(f_j(\boldsymbol{r})|z_j(\boldsymbol{r})) \\
p(A_{ij}) = \mathcal{N}(A_{0\,ij}, \sigma_{0\,ij}^2) \quad \text{or} \quad p(\text{vect}(\boldsymbol{A})) = \mathcal{N}(\text{vect}(\boldsymbol{A}_0), \sigma_{0\,ij}^2 \boldsymbol{I})
\end{array}\right.$$

$$\left\{\begin{array}{l}
p(\underline{\boldsymbol{f}}|\underline{\boldsymbol{g}}, \underline{\boldsymbol{z}}, \underline{\boldsymbol{\theta}}, \boldsymbol{A}) = \sum_{\boldsymbol{r}\in\mathcal{R}} \mathcal{N}(\widehat{\boldsymbol{f}}(\boldsymbol{r}), \widehat{\boldsymbol{\Sigma}}(\boldsymbol{r})) \quad \text{with} \\
\widehat{\boldsymbol{\Sigma}}(\boldsymbol{r}) = (\boldsymbol{A}^t {\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}}^{-1} \boldsymbol{A} + {\boldsymbol{\Sigma}_z}^{-1}(\boldsymbol{r}))^{-1} \quad \text{and} \\
\widehat{\boldsymbol{f}}(\boldsymbol{r}) = \widehat{\boldsymbol{\Sigma}}(\boldsymbol{r}) \left(\boldsymbol{A}^t {\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}}^{-1} \boldsymbol{g}(\boldsymbol{r}) + \boldsymbol{\Sigma}_z(\boldsymbol{r})^{-1} \boldsymbol{m}_z(\boldsymbol{r})\right) \\
p(\boldsymbol{z}(\boldsymbol{r})|\boldsymbol{g}(\boldsymbol{r}), \boldsymbol{\theta}, \boldsymbol{A}) \propto (p(\boldsymbol{g}(\boldsymbol{r})|\boldsymbol{z}(\boldsymbol{r}), \boldsymbol{\theta}, \boldsymbol{A}))\ p(\boldsymbol{z}(\boldsymbol{r})) \quad \text{with} \\
p(\boldsymbol{g}(\boldsymbol{r})|\boldsymbol{z}(\boldsymbol{r}), \underline{\boldsymbol{\theta}}) = \mathcal{N}(\boldsymbol{A}\boldsymbol{m}_{z(\boldsymbol{r})}, \boldsymbol{A}\boldsymbol{\Sigma}_{z(\boldsymbol{r})}\boldsymbol{A}^t + \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}) \\
p(\boldsymbol{A}, \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}|\underline{\boldsymbol{f}}, \underline{\boldsymbol{g}}, \underline{\boldsymbol{\theta}}) = \mathcal{N}(\boldsymbol{A};\ \boldsymbol{A}_p, \boldsymbol{\Sigma}_p)\ \mathcal{W}({\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}}^{-1};\ \nu_p, \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}p})
\end{array}\right.$$


Mixing
Separation
Segmentation



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

Single channel image restoration

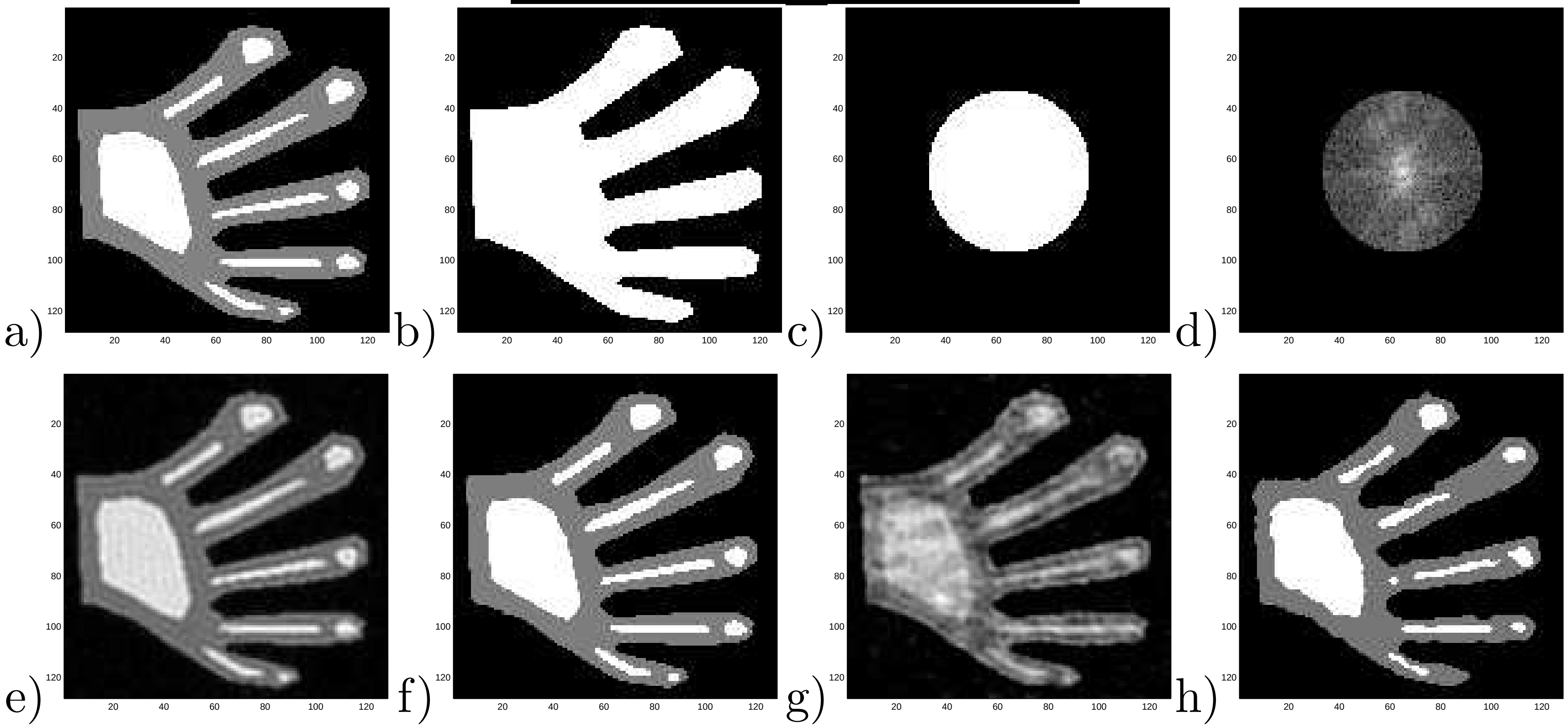
$$g(\boldsymbol{r}') = \int h(\boldsymbol{r}' - \boldsymbol{r})\, f(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r} + \epsilon(\boldsymbol{r}') \longrightarrow \boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\epsilon}$$

Fourier synthesis inverse problem

$$g(\boldsymbol{\omega}) = \int \exp\left[-j(\boldsymbol{\omega}.\boldsymbol{r})\right]\, f(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r} + \epsilon(\boldsymbol{\omega}) \longrightarrow \boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} + \boldsymbol{\epsilon}$$

$$\left\{\begin{array}{l}
p(\boldsymbol{\epsilon}) = \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}) \longrightarrow p(\boldsymbol{g}|\boldsymbol{f}) = \mathcal{N}(\boldsymbol{H}\boldsymbol{f}, \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}) \quad \text{with} \quad \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}} = {\sigma_{\epsilon}}^2 \boldsymbol{I} \\
p(f(\boldsymbol{r})|z(\boldsymbol{r}) = k) = \mathcal{N}(m_k, \sigma_k^2), \quad k = 1, \cdots, K \\
p(\boldsymbol{f}|\boldsymbol{z}, \boldsymbol{\theta}, \boldsymbol{g}) = \mathcal{N}(\widehat{\boldsymbol{f}}, \widehat{\boldsymbol{\Sigma}}) \quad \text{with} \\
\widehat{\boldsymbol{\Sigma}} = (\boldsymbol{H}^t {\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}}^{-1} \boldsymbol{H} + {\boldsymbol{\Sigma}_z}^{-1})^{-1} \quad \text{and} \quad \widehat{\boldsymbol{f}} = \widehat{\boldsymbol{\Sigma}} \left(\boldsymbol{H}^t {\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}}^{-1} \boldsymbol{g} + {\boldsymbol{\Sigma}_z}^{-1} \boldsymbol{m}_z\right) \\
\text{Compute} \quad \widehat{\boldsymbol{f}} = \arg\max_{\boldsymbol{f}} \left\{p(\boldsymbol{f}|\boldsymbol{z}, \boldsymbol{\theta}, \boldsymbol{g})\right\} = \arg\min_{\boldsymbol{f}} \left\{J(\boldsymbol{f})\right\} \quad \text{with} \\
J(\boldsymbol{f}) = \frac{1}{\sigma_\epsilon^2} \|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{f}\|^2 + \sum_k \frac{\|\boldsymbol{f}_k - m_k \boldsymbol{1}\|^2}{\sigma_k^2} \\
p(\boldsymbol{z}|\boldsymbol{g}, \boldsymbol{\theta}) \propto p(\boldsymbol{g}|\boldsymbol{z}, \boldsymbol{\theta})\, p(\boldsymbol{z}) \quad \text{with} \\
p(\boldsymbol{g}|\boldsymbol{z}, \boldsymbol{\theta}) = \mathcal{N}(\boldsymbol{H}\boldsymbol{m}_z, \boldsymbol{\Sigma}_{\boldsymbol{g}}) \quad \text{with} \quad \boldsymbol{\Sigma}_{\boldsymbol{g}} = \boldsymbol{H}\boldsymbol{\Sigma}_z \boldsymbol{H}^t + \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}} \\
\text{Use} \quad p(\boldsymbol{z}|\boldsymbol{g}, \boldsymbol{f}, \boldsymbol{\theta}) \propto p(\boldsymbol{g}|\boldsymbol{f}, \boldsymbol{z}, \boldsymbol{\theta})\, p(\boldsymbol{z})
\end{array}\right.$$

Simulation results

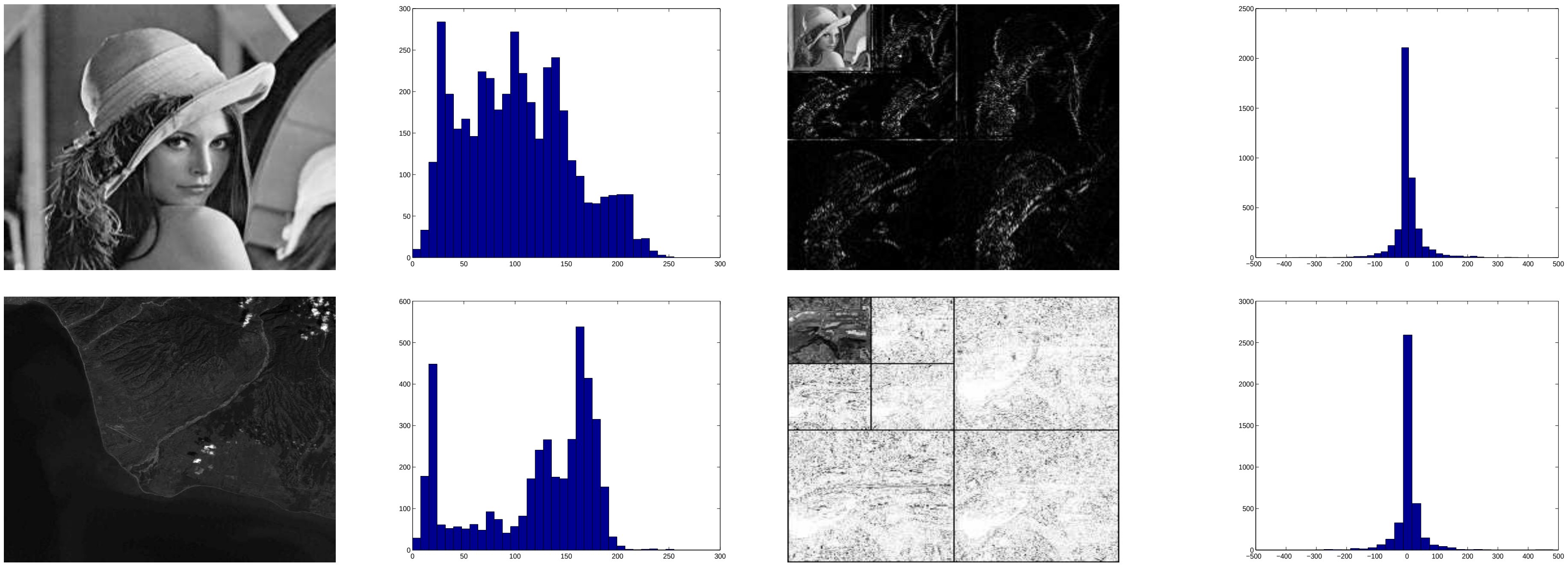


a) object, b) exact known support, c) support of the data, d) measured data, e) and f) Results when phase is measured: e) IFT and f) proposed method, g) and h) Results when the phase is not measured but we know the support of the object: g) by Gerchberg-Saxton h) by the proposed method.

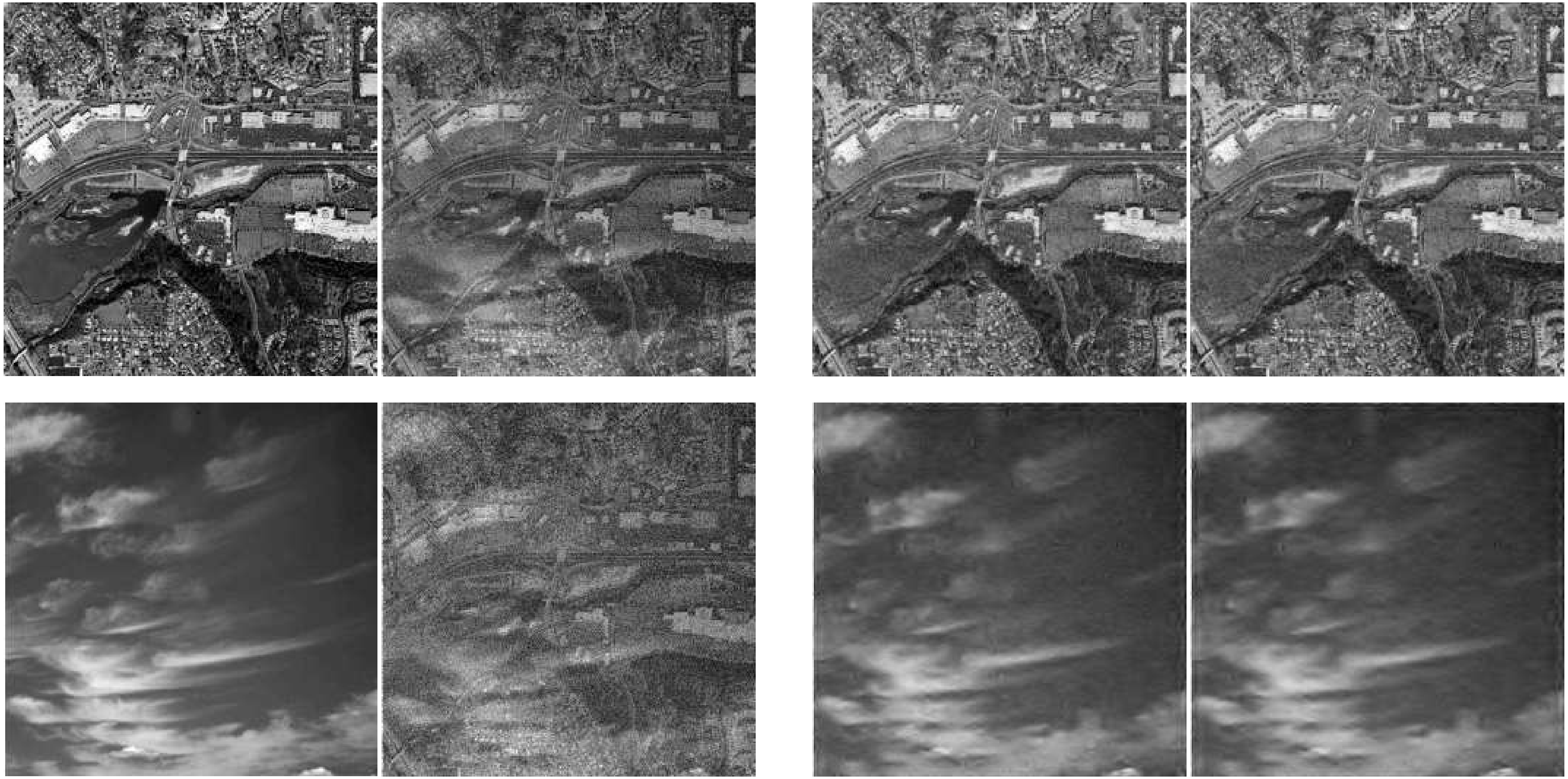
Wavelet domain Bayesian image processing

$$\boldsymbol{g}(\boldsymbol{r}) = \boldsymbol{f}(\boldsymbol{r}) + \boldsymbol{\epsilon}(\boldsymbol{r}) \longrightarrow \mathrm{WT} \longrightarrow \widetilde{\boldsymbol{g}}^j(\boldsymbol{r}) = \widetilde{\boldsymbol{f}}^j(\boldsymbol{r}) + \widetilde{\boldsymbol{\epsilon}}^j(\boldsymbol{r})$$

images | Hist. of images | Wavelet coeff. | Hist. of wavelet coeff.



- multi-resolution computation
- Wavelet coefficients can be classified and segmented in only $K = 2$ classes



$\underline{\boldsymbol{f}}$ $\underline{\boldsymbol{g}}$ $\widehat{\underline{\boldsymbol{f}}}$ (GM) $\widehat{\underline{\boldsymbol{f}}}$ (HMM)

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 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 : Super-resolution from a set of lower resolution images
- N. Bali : Source separation using different Hidden Markov models for images
- H. Akbari : Mixture models for source separation in Hyperspectral imaging