A Hidden Markov model for Bayesian fusion of multivariate signals

Olivier Féron and Ali Mohammad-Djafari

Laboratoire des Signaux et Systèmes

Unité mixte de recherche 8506 CNRS-Supélec-UPS

Supélec, Plateau de Moulon, 91192 Gif-sur-Yvette, FRANCE.

djafari@lss.supelec.fr

http://djafari.free.fr

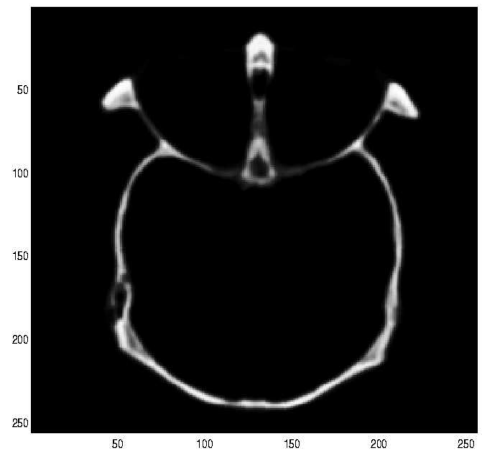
Content

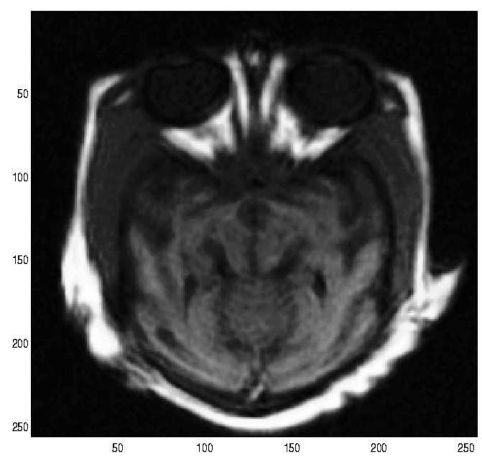
1. Examples of simple image fusion problems
2. A very simple model and basics of the Bayesian approach
3. A more realistic model and the need for hidden markov modeling
4. Expression of the posterior law and MCMC algorithm implementation
5. Simulation results
6. Conclusions

Examples of simple image fusion problems

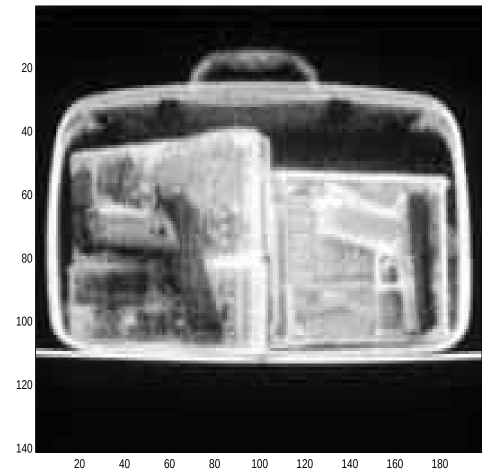
Photography

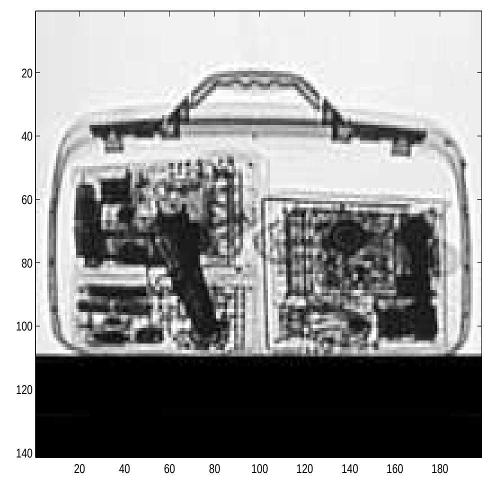
Medical imaging





Security systems





A very simple modeling

$$g_i(r) = f(r) + \epsilon_i(r), \quad r = (x, y), \quad i = 1, 2 \quad \longrightarrow \quad \boldsymbol{g}_i = \boldsymbol{f} + \boldsymbol{\epsilon}_i$$

- Hypothesis on $\epsilon_i \longrightarrow \qquad p(\boldsymbol{g}_i|\boldsymbol{f})$
- Hypothesis on $\boldsymbol{f} \longrightarrow \qquad p(\boldsymbol{f})$
- Bayes rule: $p(\boldsymbol{f}|\boldsymbol{g}_1, \boldsymbol{g}_2)$
- Bayesian MAP estimation:

$$\widehat{\boldsymbol{f}} = \arg\max_{\boldsymbol{f}} \{p(\boldsymbol{f}|\boldsymbol{g}_1, \boldsymbol{g}_2)\} = \arg\min_{\boldsymbol{f}} \{J(\boldsymbol{f}) = -\ln p(\boldsymbol{f}|\boldsymbol{g}_1, \boldsymbol{g}_2)\}$$

— Case of Generalized Exponential priors:

$$p(\boldsymbol{g}_i|\boldsymbol{f}_i) = p_{\boldsymbol{\epsilon}_i}(\boldsymbol{g}_i - \boldsymbol{f}_i) \propto \exp\{-\frac{1}{\sigma_i^2}\|\boldsymbol{g}_i - \boldsymbol{f}_i\|^\beta\}, \quad k = 1, 2$$

$$p(\boldsymbol{f}_i) \propto \exp\{-\frac{1}{\sigma_0^2}\|\boldsymbol{f} - \boldsymbol{f}_0\|^\alpha\}$$

$$J(\boldsymbol{f}) = \sum_i \frac{1}{\sigma_i^2} \|\boldsymbol{g}_i - \boldsymbol{f}\|^\beta + \frac{1}{\sigma_0^2} \|\boldsymbol{f} - \boldsymbol{f}_0\|^\alpha \quad \text{with} \quad \|\boldsymbol{f}\|^\alpha = \iint |f(x,y)|^\alpha \,\mathrm{d}x \,\mathrm{d}y$$

— Case of Gaussian priors ($\alpha = \beta = 2$):

$$\nabla J = 0 \longrightarrow \widehat{\boldsymbol{f}} = \frac{1}{\sum_i \lambda_i} \left(\sum_i \lambda_i \boldsymbol{g}_i + \lambda_0 \boldsymbol{f}_0 \right), \quad \lambda_i = \frac{1}{\sigma_i^2}, \quad i = 0, \ldots, M$$

$$\lambda_0 = 0, \lambda_1 = \ldots = \lambda_M = 1 \longrightarrow \widehat{\boldsymbol{f}} = \frac{1}{M} \sum_i \boldsymbol{g}_i \quad \text{mean}$$

— Case of Exponential ($\alpha = 0, \beta = 1$): $\widehat{\boldsymbol{f}}$ =median$\{\boldsymbol{g}_i, i = 1, \ldots, M\}$

A more realistic modeling

$$g_i(r) = f_i(r) + \epsilon_i(r), \quad i = 1,2 \quad \longrightarrow \quad \boldsymbol{g}_i = \boldsymbol{f}_i + \boldsymbol{\epsilon}_i$$

$$p(\boldsymbol{g}_i|\boldsymbol{f}_i) = p_{\boldsymbol{\epsilon}_i}(\boldsymbol{g}_i - \boldsymbol{f}_i) = \mathcal{N}(\boldsymbol{g}_i - \boldsymbol{f}_i, \Sigma_{\boldsymbol{\epsilon}_i}), \quad k = 1,2$$

$$\boldsymbol{f} = \begin{pmatrix} \boldsymbol{f}_1 \\ \boldsymbol{f}_2 \end{pmatrix} \sim \mathcal{N}\left(\boldsymbol{f}_0 = \begin{pmatrix} \boldsymbol{f}_{1_0} \\ \boldsymbol{f}_{2_0} \end{pmatrix}, \boldsymbol{\Sigma}_{\boldsymbol{f}} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}\right)$$

$$p(\boldsymbol{f}|\boldsymbol{g}) = \mathcal{N}(\widehat{\boldsymbol{f}}, \widehat{\boldsymbol{\Sigma}}) \text{ with } \widehat{\boldsymbol{\Sigma}} = (\boldsymbol{H}^t \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}} \boldsymbol{H} + \boldsymbol{\Sigma}_{\boldsymbol{f}})^{-1} \text{ and } \widehat{\boldsymbol{f}} = \widehat{\boldsymbol{f}}_0 + \widehat{\boldsymbol{\Sigma}} \boldsymbol{H}^t (\boldsymbol{g} - \boldsymbol{H} \boldsymbol{f}_0)$$

Remark :

When $\boldsymbol{f}_1$ and $\boldsymbol{f}_2$ are assumed independent $\boldsymbol{\Sigma}_{12} = \boldsymbol{\Sigma}_{21} = 0 \longrightarrow$

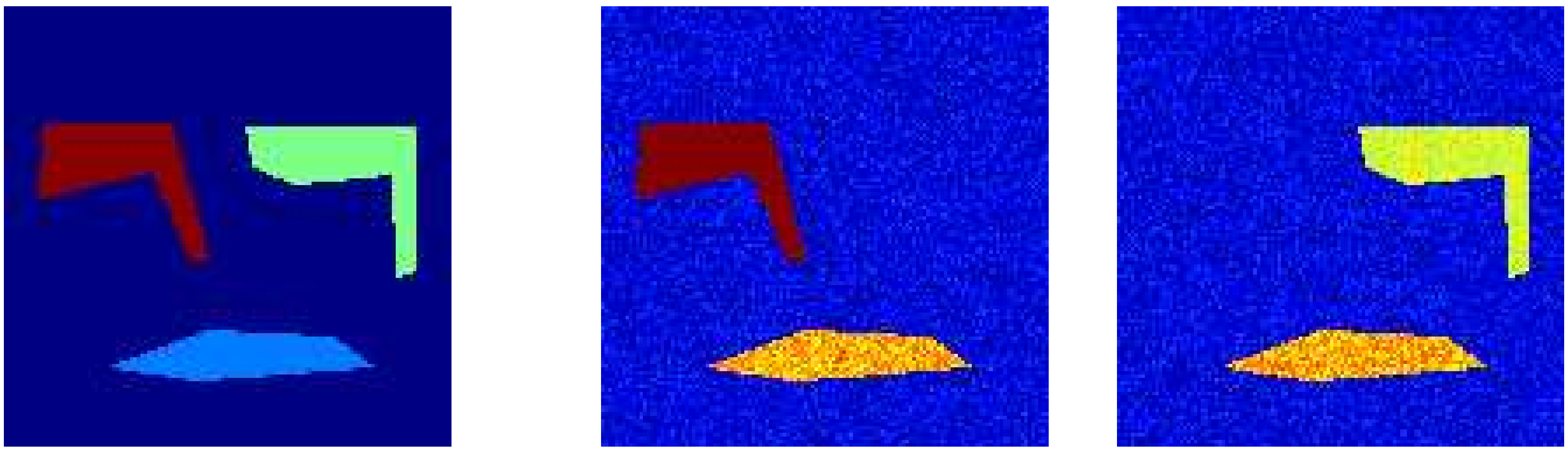
$$p(\boldsymbol{f}_i|\boldsymbol{g}_i) = \mathcal{N}(\widehat{\boldsymbol{f}}_i, \widehat{\boldsymbol{\Sigma}}_i)$$

$$\text{avec } \widehat{\boldsymbol{\Sigma}}_i = (\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}_i} + \boldsymbol{\Sigma}_{\boldsymbol{f}_i})^{-1} \text{ and } \widehat{\boldsymbol{f}}_i = \widehat{\boldsymbol{f}}_{0_i} + \widehat{\boldsymbol{\Sigma}}_i (\boldsymbol{g}_i - \boldsymbol{f}_{0_i})$$

Hidden Markov Modeling

$$g_i(r) = f_i(r) + \epsilon_i(r), \quad i = 1, 2 \quad \longrightarrow \quad \boldsymbol{g}_i = \boldsymbol{f}_i + \boldsymbol{\epsilon}_i$$

- The only thing the two images share is their "anatomy" $z(r)$



- Regions in the image with same materials (homogeneous regions):

$$R_k = \{r : z(r) = k\}, \qquad |R_k| = n_k \longrightarrow \boldsymbol{f}_{ik} = \{f_i(r) : z(r) = k\}$$

- $z(r)$ or equivalently $\boldsymbol{z}$ is a hidden variable common to the images $\longrightarrow p(\boldsymbol{f}_i|\boldsymbol{z})$

- Modeling the homogeneity:

$$p(f_i(r)|z(r)=k) = \mathcal{N}(m_{i\,k}, \sigma_{i\,k}^2) \quad \longrightarrow \quad p(\boldsymbol{f}_{i\,k}) = \mathcal{N}(m_{i\,k}\mathbf{1}, \sigma_{i\,k}^2 \boldsymbol{I})$$

$$\begin{aligned} p(\boldsymbol{f}_i|\boldsymbol{z}) &= \prod_{k=1}^{K} \mathcal{N}(m_{i\,k}\mathbf{1}, \sigma_{i\,k}^2 \boldsymbol{I}) \\ &= \prod_{k=1}^{K} \left(\frac{1}{\sqrt{2\pi\sigma_{i\,k}^2}}\right)^{n_k} \exp\left\{-\frac{1}{2\sigma_{i\,k}^2} ||\boldsymbol{f}_{i\,k} - m_{i\,k}\mathbf{1}||^2\right\} \end{aligned}$$

- $\boldsymbol{\epsilon}_i$ centered, Gaussian, i.i.d. :

$$p(\boldsymbol{g}_1, \boldsymbol{g}_2 | \boldsymbol{f}_1, \boldsymbol{f}_2) = \prod_{i=1}^{2} p(\boldsymbol{g}_i | \boldsymbol{f}_i) = \prod_{i=1}^{2} p_{\epsilon_i}(\boldsymbol{g}_i - \boldsymbol{f}_i) = \prod_{i=1}^{2} \mathcal{N}(\boldsymbol{g}_i - \boldsymbol{f}_i, \sigma_{\epsilon_i}^2 \boldsymbol{I})$$

- Markov Random Field model for $z(r)$:

$$p(\boldsymbol{z}|\alpha) = \frac{1}{T(\alpha)} \exp\left\{ \alpha \sum_{r \in \mathcal{S}} \sum_{s \in \mathcal{V}(r)} \delta(z(r) - z(s)) \right\}$$

- Joint posterior law:

$$\begin{aligned} p(\boldsymbol{f}_1, \boldsymbol{f}_2, \boldsymbol{z} | \boldsymbol{g}_1, \boldsymbol{g}_2) &= p(\boldsymbol{f}_1, \boldsymbol{f}_2 | \boldsymbol{z}, \boldsymbol{g}_1, \boldsymbol{g}_2)\, p(\boldsymbol{z} | \boldsymbol{g}_1, \boldsymbol{g}_2) \\ &\propto \prod_{i=1}^{2} p(\boldsymbol{g}_i | \boldsymbol{f}_i)\, p(\boldsymbol{f}_i | \boldsymbol{z})\, p(\boldsymbol{g}_i | \boldsymbol{z}) \quad p(\boldsymbol{z}) \end{aligned}$$

- Hyperparameters: $\boldsymbol{\theta}_i = \left\{\sigma_{\epsilon_i}, \{m_{ik}, \sigma_{ik}^2\}\right\}$
- Conjugate priors:

$$
\begin{aligned}
p(\sigma_0) &= \mathcal{IG}(\alpha_0, \beta_0) \\
p(m_{ik}) &= \mathcal{N}(m_{i0}, \sigma_{i\,0}^2) \\
p(\sigma_{ik}) &= \mathcal{IG}(\alpha_{i0}, \beta_{i0})
\end{aligned}
$$

- Joint posterior law:

$$
\begin{aligned}
p(\boldsymbol{f}_1, \boldsymbol{f}_2, \boldsymbol{z}, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2 | \boldsymbol{g}_1, \boldsymbol{g}_2) &= p(\boldsymbol{f}_1, \boldsymbol{f}_2 | \boldsymbol{z}, \boldsymbol{g}_1, \boldsymbol{g}_2, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2) p(\boldsymbol{z} | \boldsymbol{g}_1, \boldsymbol{g}_2, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2) p(\boldsymbol{\theta}_1) p(\boldsymbol{\theta}_2) \\
&\propto \prod_{i=1}^{2} \quad p(\boldsymbol{g}_i | \boldsymbol{f}_i, \sigma_{\epsilon_i})\, p(\boldsymbol{f}_i | \boldsymbol{z}, \{m_{ik}, \sigma_{ik}^2\}) \\
&\qquad \times p(\boldsymbol{g}_i | \boldsymbol{z}, \{m_{ik}, \sigma_{ik}^2\}) \quad p(\boldsymbol{\theta}_i) \quad p(\boldsymbol{z} | \alpha)
\end{aligned}
$$

$$
\begin{aligned}
p(\boldsymbol{f}_1, \boldsymbol{f}_2, \boldsymbol{z}, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2 | \boldsymbol{g}_1, \boldsymbol{g}_2) &= p(\boldsymbol{f}_1, \boldsymbol{f}_2, \boldsymbol{z} | \boldsymbol{g}_1, \boldsymbol{g}_2, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2)\, p(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 | \boldsymbol{g}_1, \boldsymbol{g}_2) \\
&\propto p(\boldsymbol{z}|\alpha) \prod_{i=1}^{2} p(\boldsymbol{f}_i | \boldsymbol{g}_i, \boldsymbol{z}, \boldsymbol{\theta}_i)\, p(\boldsymbol{\theta}_i)
\end{aligned}
$$

MCMC Gibbs sampling:

---

repeat until convergence

1. simulate $(\hat{\boldsymbol{f}}_1^{(n)}, \hat{\boldsymbol{f}}_2^{(n)}, \hat{\boldsymbol{z}}^{(n)}) \sim p(\boldsymbol{f}_1, \boldsymbol{f}_2, \boldsymbol{z} | \boldsymbol{g}_1, \boldsymbol{g}_2, \hat{\boldsymbol{\theta}}_1^{(n-1)}, \hat{\boldsymbol{\theta}}_2^{(n-1)})$
2. simulate $\hat{\boldsymbol{\theta}}_i^{(n)} \sim p(\boldsymbol{\theta}_i | \boldsymbol{g}_i, \hat{\boldsymbol{f}}_i^{(n)}, \hat{\boldsymbol{z}}^{(n)})$

Gibbs sampling using the Markovian structure of $\boldsymbol{z}$:

---

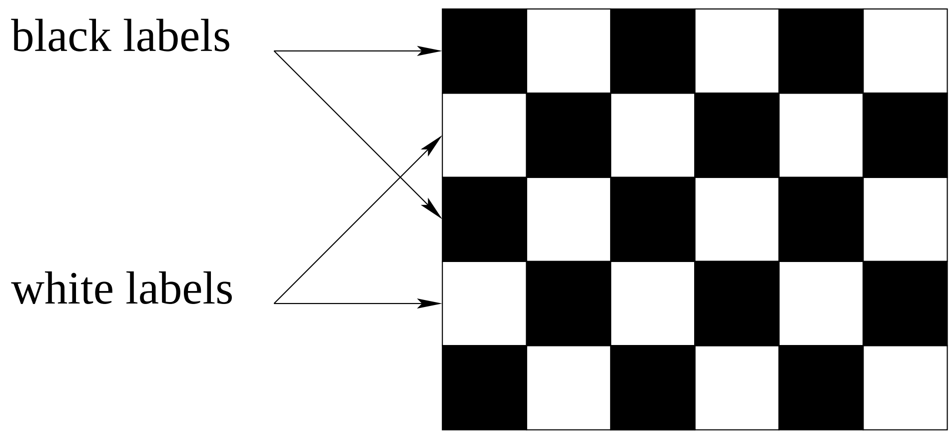
repeat until convergence

1. simulate $\hat{\boldsymbol{z}}^{(n)} \sim p(\boldsymbol{z} | \hat{\boldsymbol{z}}^{(n-1)}, \boldsymbol{g}_1, \boldsymbol{g}_2, \hat{\boldsymbol{\theta}}_1^{(n-1)}, \hat{\boldsymbol{\theta}}_2^{(n-1)})$

   simulate $\hat{\boldsymbol{f}}_i^{(n)} \sim p(\boldsymbol{f}_i | \boldsymbol{g}_i, \hat{\boldsymbol{z}}^{(n)}, \hat{\boldsymbol{\theta}}_i^{(n-1)})$
2. simulate $\hat{\boldsymbol{\theta}}_i^{(n)} \sim p(\boldsymbol{\theta}_i | \hat{\boldsymbol{f}}_i^{(n)}, \hat{\boldsymbol{z}}^{(n)}, \boldsymbol{g}_i)$

Parallel implementation of the Gibbs sampling

Chessboard decomposition of the labels z



repeat until convergence

1. simulate $\hat{\boldsymbol{z}}_W^{(n)} \sim p(\boldsymbol{z}|\hat{\boldsymbol{z}}_B^{(n-1)}, \boldsymbol{g}_1, \boldsymbol{g}_2, \hat{\boldsymbol{\theta}}_1^{(n-1)}, \hat{\boldsymbol{\theta}}_2^{(n-1)})$

   simulate $\hat{\boldsymbol{z}}_B^{(n)} \sim p(\boldsymbol{z}|\hat{\boldsymbol{z}}_W^{(n)}, \boldsymbol{g}_1, \boldsymbol{g}_2, \hat{\boldsymbol{\theta}}_1^{(n-1)}, \hat{\boldsymbol{\theta}}_2^{(n-1)})$

   simulate $\hat{\boldsymbol{f}}_i^{(n)} \sim p(\boldsymbol{f}_i|\boldsymbol{g}_i, \hat{\boldsymbol{z}}^{(n)}, \hat{\boldsymbol{\theta}}_i^{(n-1)})$
2. simulate $\hat{\boldsymbol{\theta}}_i^{(n)} \sim p(\boldsymbol{\theta}_i|\hat{\boldsymbol{f}}_i^{(n)}, \hat{\boldsymbol{z}}^{(n)}, \boldsymbol{g}_i)$

Sampling $\boldsymbol{f}_1, \boldsymbol{f}_2, \boldsymbol{z} | \boldsymbol{g}_1, \boldsymbol{g}_2, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2$:

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

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

$$p(f_i(r) | g_i(r), z(r) = k, \boldsymbol{\theta}_i) = \mathcal{N}(m_{i\,k}^{apost}, \sigma_{i\,k}^{2\,apost})$$

where

$$\begin{aligned} \sigma_{i\,k}^{2\,apost} &= \left(\frac{1}{\sigma_{\varepsilon_i}^2} + \frac{1}{\sigma_{i\ k}^2} \right)^{-1} \\ m_{i\,k}^{apost} &= \sigma_{i\,k}^{2\,apost} \left(\frac{g_i(r)}{\sigma_{\varepsilon_i}^2} + \frac{m_{i\,k}}{\sigma_{i\ k}^2} \right). \end{aligned}$$

Sampling $\boldsymbol{\theta}_i|\boldsymbol{f}_i, \boldsymbol{g}_i, \boldsymbol{z}$:

$$p(\boldsymbol{\theta}_i|\boldsymbol{f}_i, \boldsymbol{g}_i, \boldsymbol{z}) \propto p(\sigma^2_{\varepsilon_i}|\boldsymbol{f}_i, \boldsymbol{g}_i)\; p(\boldsymbol{m}_i, \boldsymbol{\sigma}^2_i|\boldsymbol{f}_i, \boldsymbol{z})$$

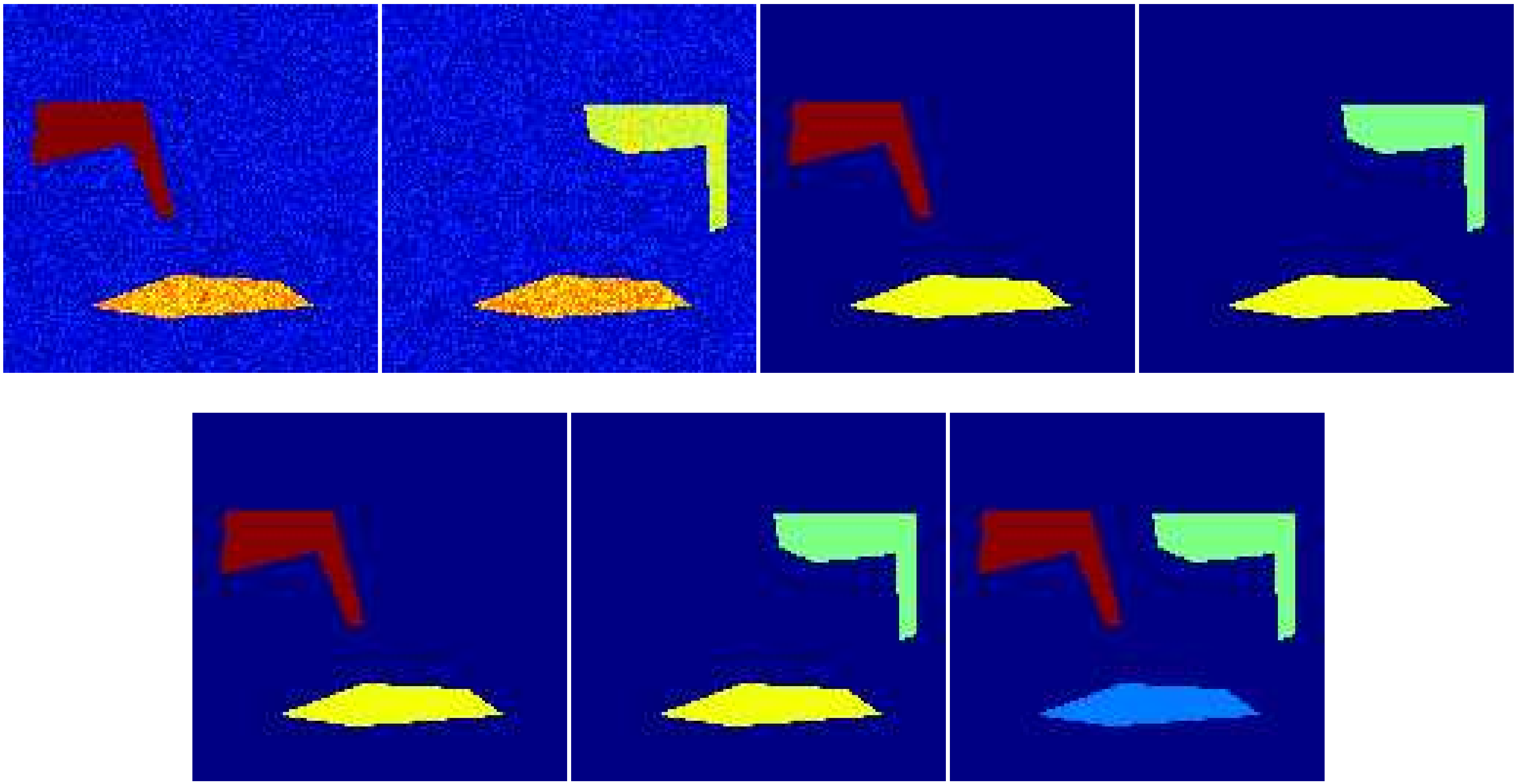
- $m_{i\,k}|\boldsymbol{f}_i, \boldsymbol{z}, \sigma^2_{i\,k}, m_{i0}, \sigma^2_{i\,0} \sim \mathcal{N}(\mu_{i\,k}, v^2_{i\,k})$ with

$$\mu_{i\,k} = v^2_{i\,k} \left(\frac{m_{i0}}{\sigma^2_{i\,0}} + \frac{1}{\sigma^2_{i\,k}} \sum_{r\in R_k} f_i(r) \right) \quad \text{and} \quad v^2_{i\,k} = \left(\frac{n_k}{\sigma^2_{i\,k}} + \frac{1}{\sigma^2_{i\,0}} \right)^{-1}$$

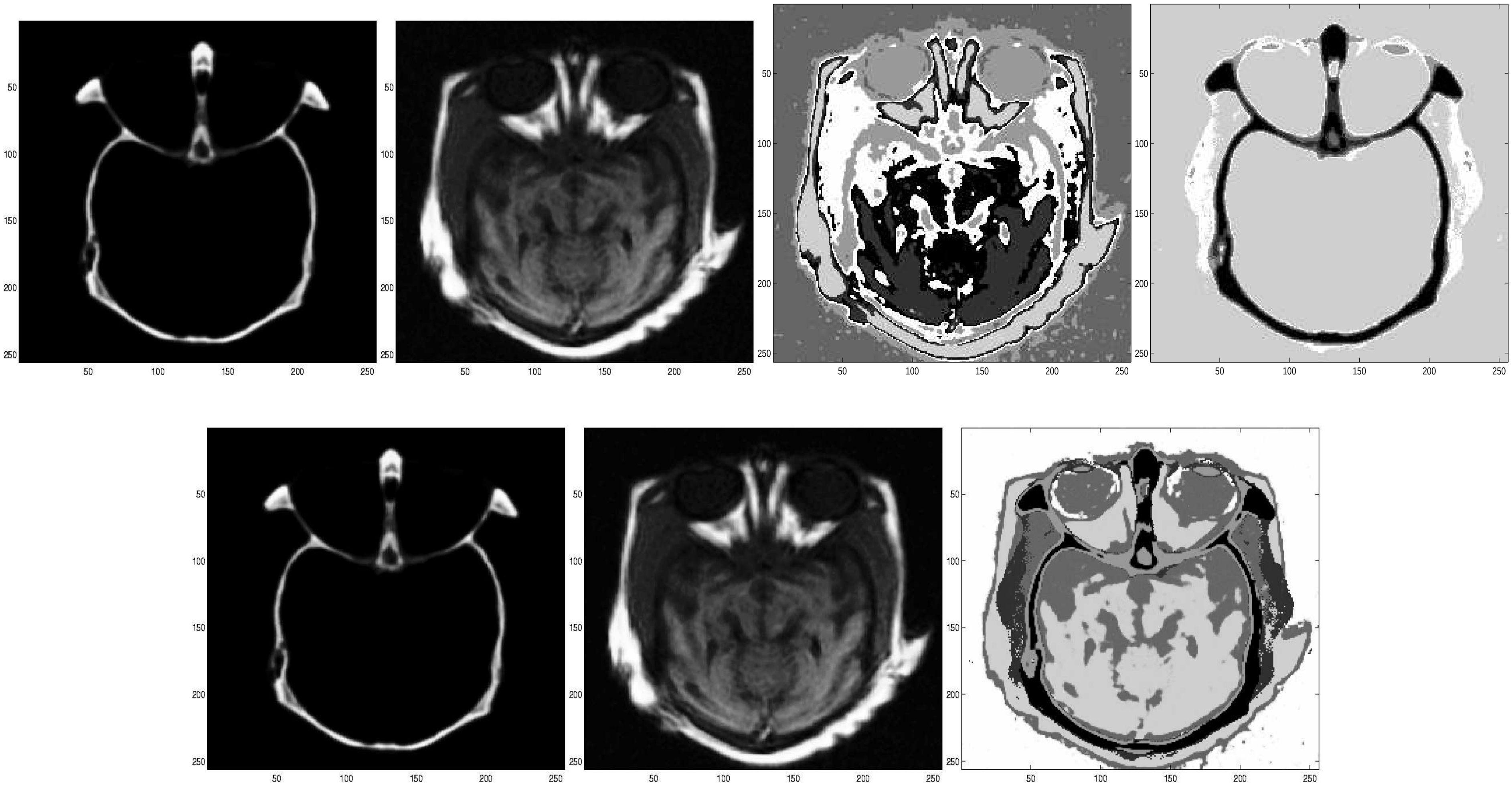
- $\sigma^2_{i\,k}|\boldsymbol{f}_i, \boldsymbol{z}, \alpha_{i0}, \beta_{i0} \sim \mathcal{IG}(\alpha_{i\,k}, \beta_{i\,k})$ with $\alpha_{i\,k} = \alpha_{i0} + \frac{n_k}{2}$ and $\beta_{i\,k} = \beta_{i0} + \frac{s_i}{2}$ where $s_i = \sum_{r\in R_k} \left(f_i(r)^2 - n_k \bar{f_i}^2 \right)$ and $\bar{f_i} = \frac{1}{n_k} \sum_{r\in R_k} f_i(r)$

- $\sigma^2_{\varepsilon_i}|\boldsymbol{f}_i, \boldsymbol{g}_i \sim \mathcal{IG}(\nu_i, \Sigma_i)$ with $\nu_i = \frac{S}{2}$ and $\Sigma_i = \frac{S}{2} \left(R_{g_i g_i} - \frac{R^2_{g_i f_i}}{f_i f_i} \right)$, $\qquad S =$ total number of pixels where $R_{xy} = \frac{1}{S} \sum_r x(r) y^*(r)$.

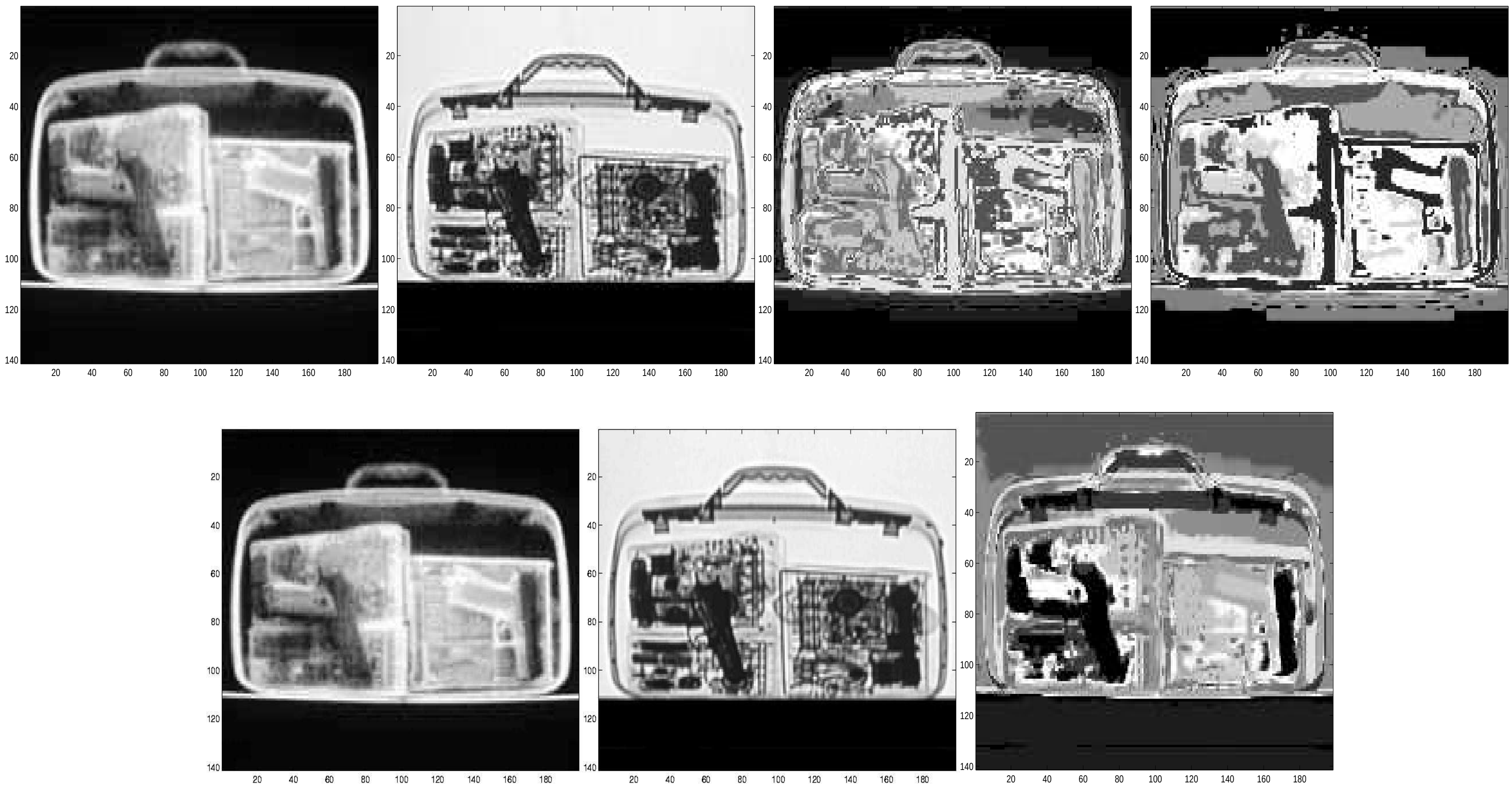
Simulation results



Data fusion in medical imaging



Data fusion in Security systems



Conclusions

- We proposed a Hidden Markov Modeling for images to be used in multivariate data fusion problems.
- The Bayesian approach is the most appropriate approach for those data fusion problems.
- In a first step, we assumed that the two images are registered, but estimation of the parameters of registration can very easily be accounted for.
- In the proposed method, we assumed that all the pixels inside a region are iid. We are working to an extension of the proposed model for accounting the local correlation of pixels inside each region.