



Bayesian Blind Deconvolution of sparse images with a Student-t a priori model

Ali Mohammad-Djafari

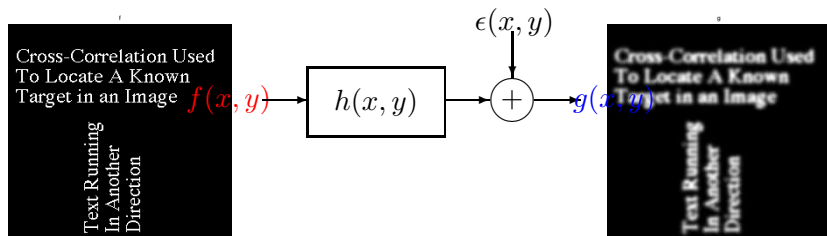
Laboratoire des Signaux et Systèmes,
UMR8506 CNRS-SUPELEC-UNIV PARIS SUD 11
SUPELEC, 91192 Gif-sur-Yvette, France
<http://lss.supelec.free.fr>

Email: djafari@lss.supelec.fr
<http://djafari.free.fr>

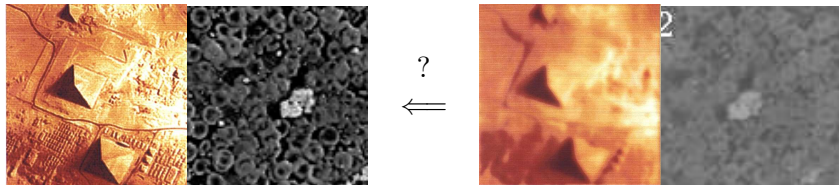
Image Restoration

Forward model: 2D Convolution

$$g(x, y) = \iint f(x', y') h(x - x', y - y') dx' dy' + \epsilon(x, y)$$



Inversion: Deconvolution



Blind Image Deconvolution

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

- ▶ Convolution: Given f and h compute g
- ▶ Identification: Given f and g estimate h
- ▶ Deconvolution: Given g and h estimate f
- ▶ Blind deconvolution: Given g estimate both h and f

Discretization:

- ▶ $g = h * f + \epsilon \longrightarrow g = \mathbf{H}f + \epsilon$,
 \mathbf{H} huge dimensional TBT matrix
obtained from the elements of Point Spread Function (PSF)
 $h(x, y)$
- ▶ $g = f * h + \epsilon \longrightarrow g = \mathbf{F}h + \epsilon$,
 \mathbf{F} huge dimensional HBT matrix
obtained from the elements of the input image $f(x, y)$

Identification and Deconvolution: Classical methods

Deconvolution

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

Wiener filtering

$$\hat{f}(u, v) = \frac{H^*(u, v)}{|H(u, v)|^2 + \frac{S_{\epsilon\epsilon}(u, v)}{S_{ff}(u, v)}} g(u, v)$$

Regularization

$$g = H f + \epsilon$$

$$J(f) = \|g - H f\|^2 + \lambda_f \|D_f f\|^2$$

$$\nabla J(f) = -2H'(g - H f) + 2\lambda_f D_f' D_f f$$

$$\hat{f} = [H'H + \lambda_f D_f' D_f]^{-1} H' g$$

Circulante approximation

$$\hat{f}(u, v) = \frac{H^*(u, v)}{|H(u, v)|^2 + \lambda_f |D_f(u, v)|^2} g(u, v)$$

Identification

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

Wiener filtering

$$\hat{h}(u, v) = \frac{F^*(u, v)}{|F(u, v)|^2 + \frac{S_{\epsilon\epsilon}(u, v)}{S_{hh}(u, v)}} g(u, v)$$

Regularization

$$g = F h + \epsilon$$

$$J(h) = \|g - F h\|^2 + \lambda_h \|D_h h\|^2$$

$$\nabla J(h) = -2F'(g - F h) + 2\lambda_h D_h' D_h h$$

$$\hat{h} = [F'F + \lambda_h D_h' D_h]^{-1} F' g$$

Circulante approximation

$$\hat{h}(u, v) = \frac{F^*(u, v)}{|F(u, v)|^2 + \lambda_h |D_h(u, v)|^2} g(u, v)$$

Deconvolution, Identification: Bayesian approach

	Deconvolution	Identification
Forward model	$\mathbf{g} = \mathbf{H} \mathbf{f} + \epsilon$	$\mathbf{g} = \mathbf{F} \mathbf{h} + \epsilon$
Likelihood	$p(\mathbf{g} \mathbf{f}) = \mathcal{N}(\mathbf{g} \mathbf{H} \mathbf{f}, \Sigma_\epsilon)$	$p(\mathbf{g} \mathbf{h}) = \mathcal{N}(\mathbf{g} \mathbf{F} \mathbf{h}, \Sigma_\epsilon)$
A priori	$p(\mathbf{f}) = \mathcal{N}(\mathbf{f} 0, \Sigma_f)$	$p(\mathbf{h}) = \mathcal{N}(\mathbf{h} 0, \Sigma_h)$
Bayes	$p(\mathbf{f} \mathbf{g}) = \frac{p(\mathbf{g} \mathbf{f})p(\mathbf{f})}{p(\mathbf{g})}$	$p(\mathbf{h} \mathbf{g}) = \frac{p(\mathbf{g} \mathbf{h})p(\mathbf{h})}{p(\mathbf{g})}$

Particular case: $\Sigma_\epsilon = \sigma_\epsilon^2 \mathbf{I}$, $\Sigma_f = \sigma_f^2 (\mathbf{D}'_f \mathbf{D}_f)^{-1}$, $\Sigma_h = \sigma_h^2 (\mathbf{D}'_h \mathbf{D}_h)^{-1}$

Deconvolution	Identification
$p(\mathbf{f} \mathbf{g}) \propto p(\mathbf{g} \mathbf{f}) p(\mathbf{f}) \propto \exp \{-J(\mathbf{f})\}$	$p(\mathbf{h} \mathbf{g}) \propto p(\mathbf{g} \mathbf{h}) p(\mathbf{h}) \propto \exp \{-J(\mathbf{h})\}$
$J(\mathbf{f}) = \ \mathbf{g} - \mathbf{H} \mathbf{f}\ ^2 + \lambda_f \ \mathbf{D}_f \mathbf{f}\ ^2$	$J(\mathbf{h}) = \ \mathbf{g} - \mathbf{F} \mathbf{h}\ ^2 + \lambda_h \ \mathbf{D}_h \mathbf{h}\ ^2$
$\lambda_f = \sigma_\epsilon^2 / \sigma_f^2$	$\lambda_h = \sigma_\epsilon^2 / \sigma_h^2$
$p(\mathbf{f} \mathbf{g}) = \mathcal{N}(\mathbf{f} \hat{\mathbf{f}}, \hat{\Sigma}_f)$	$p(\mathbf{h} \mathbf{g}) = \mathcal{N}(\mathbf{h} \hat{\mathbf{h}}, \hat{\Sigma}_h)$
$\hat{\mathbf{f}} = [\mathbf{H}' \mathbf{H} + \lambda_f \mathbf{D}'_f \mathbf{D}_f]^{-1} \mathbf{F}' \mathbf{g}$	$\hat{\mathbf{h}} = [\mathbf{F}' \mathbf{F} + \lambda_h \mathbf{D}'_h \mathbf{D}_h]^{-1} \mathbf{F}' \mathbf{g}$
$\hat{\Sigma}_f = [\mathbf{H}' \mathbf{H} + \lambda_f \mathbf{D}'_f \mathbf{D}_f]^{-1}$	$\hat{\Sigma}_h = [\mathbf{F}' \mathbf{F} + \lambda_h \mathbf{D}'_h \mathbf{D}_h]^{-1}$

Blind Deconvolution: Bayesian approach

$$\mathbf{g} = \mathbf{H} \mathbf{f} + \boldsymbol{\epsilon} = \mathbf{F} \mathbf{h} + \boldsymbol{\epsilon}$$

- ▶ Joint posterior law:

$$p(\mathbf{f}, \mathbf{h} | \mathbf{g}) \propto p(\mathbf{g} | \mathbf{f}, \mathbf{h}) p(\mathbf{f}) p(\mathbf{h}) \propto \exp \{-J(\mathbf{f}, \mathbf{h})\}$$

$$J(\mathbf{f}, \mathbf{h}) = \|\mathbf{g} - \mathbf{H} \mathbf{f}\|^2 + \lambda_f \|\mathbf{D}_f \mathbf{f}\|^2 + \lambda_h \|\mathbf{D}_h \mathbf{h}\|^2$$

- ▶ Joint MAP:

$$(\hat{\mathbf{f}}, \hat{\mathbf{h}}) = \arg \max_{(\mathbf{f}, \mathbf{h})} \{p(\mathbf{f}, \mathbf{h} | \mathbf{g})\} = \arg \min_{(\mathbf{f}, \mathbf{h})} \{J(\mathbf{f}, \mathbf{h})\}$$

- ▶ Alternate optimization:

$$\begin{cases} \hat{\mathbf{f}}^{(k)} = \arg \max_{\mathbf{f}} \{p(\mathbf{f}, \mathbf{h}^{(k)} | \mathbf{g})\} = \arg \min_{\mathbf{f}} \{J(\mathbf{f}, \mathbf{h}^{(k)})\} \\ \hat{\mathbf{h}}^{(k)} = \arg \max_{\mathbf{h}} \{p(\mathbf{f}^{(k)}, \mathbf{h} | \mathbf{g})\} = \arg \min_{\mathbf{h}} \{J(\mathbf{f}^{(k)}, \mathbf{h})\} \end{cases}$$

Blind Deconvolution: Variational Bayesian Approximation algorithm

- ▶ Joint posterior law:

$$p(\mathbf{f}, \mathbf{h} | \mathbf{g}) \propto p(\mathbf{g} | \mathbf{f}, \mathbf{h}) p(\mathbf{f}) p(\mathbf{h})$$

- ▶ Approximation: $p(\mathbf{f}, \mathbf{h} | \mathbf{g})$ by $q(\mathbf{f}, \mathbf{h}) = q_1(\mathbf{f}) q_2(\mathbf{h})$
- ▶ Criterion of approximation: Kullback-Leiler

$$\text{KL}(q|p) = \int q \ln \frac{q}{p} = \int q_1 q_2 \ln \frac{q_1 q_2}{p}$$

$$\begin{aligned} \text{KL}(q_1 q_2 | p) &= \int q_1 \ln q_1 + \int q_2 \ln q_2 - \int q \ln p \\ &= -\mathcal{H}(q_1) - \mathcal{H}(q_2) + \langle -\ln p(\mathbf{f}, \mathbf{h} | \mathbf{g}) \rangle_q \end{aligned}$$

- ▶ When the expression of q_1 and q_2 are obtained, use them.

Variational Bayesian Approximation algorithm

- ▶ Kullback-Leibler criterion

$$\begin{aligned}\text{KL}(q_1 q_2 | p) &= \int q_1 \ln q_1 + \int q_2 \ln q_2 + \int q \ln p \\ &= -\mathcal{H}(q_1) - \mathcal{H}(q_2) + \langle -\ln p(\mathbf{f}, \mathbf{h} | \mathbf{g}) \rangle_q\end{aligned}$$

- ▶ Free energy

$$\mathcal{F}(q_1 q_2) = - \langle \ln p(\mathbf{f}, \mathbf{h} | \mathbf{g}) \rangle_{q_1 q_2}$$

- ▶ Equivalence between optimization of $\text{KL}(q_1 q_2 | p)$ and $\mathcal{F}(q_1 q_2)$
- ▶ Alternate optimization:

$$\hat{q}_1 = \arg \min_{q_1} \{ \text{KL}(q_1 q_2 | p) \} = \arg \min_{q_1} \{ \mathcal{F}(q_1 q_2) \}$$

$$\hat{q}_2 = \arg \min_{q_2} \{ \text{KL}(q_1 q_2 | p) \} = \arg \min_{q_2} \{ \mathcal{F}(q_1 q_2) \}$$

Summary of Bayesian methods for Blind Deconvolution

$$p(\mathbf{f}, \mathbf{h}|\mathbf{g}) = \frac{p(\mathbf{g}|\mathbf{f}, \mathbf{h}) p(\mathbf{f}) p(\mathbf{h})}{p(\mathbf{g})}$$

- ▶ JMAP: $(\hat{\mathbf{f}}, \hat{\mathbf{h}}) = \arg \max_{(\mathbf{f}, \mathbf{h})} \{p(\mathbf{f}, \mathbf{h}|\mathbf{g})\}$

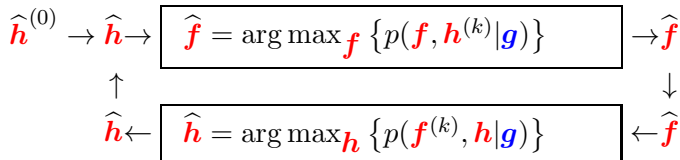
$$\begin{cases} \hat{\mathbf{f}}^{(k+1)} = \arg \max_{\mathbf{f}} \{p(\mathbf{f}, \mathbf{h}^{(k)}|\mathbf{g})\} \\ \hat{\mathbf{h}}^{(k+1)} = \arg \max_{\mathbf{h}} \{p(\mathbf{f}^{(k)}, \mathbf{h}|\mathbf{g})\} \end{cases}$$

- ▶ VBA: Approximation: $p(\mathbf{f}, \mathbf{h}|\mathbf{g})$ by $q(\mathbf{f}, \mathbf{h}) = q_1(\mathbf{f}) q_2(\mathbf{h})$

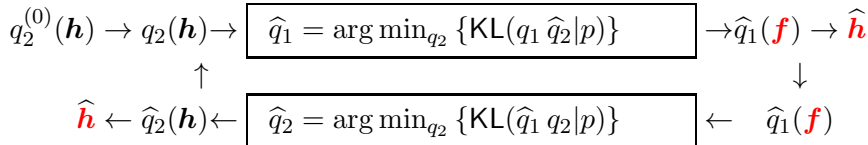
$$\begin{cases} \hat{q}_1(\mathbf{f}) = \arg \min_{q_1} \{\text{KL}(q_1 q_2|p)\} = \arg \min_{q_1} \{\mathcal{F}(q_1 q_2)\} \\ \hat{q}_2(\mathbf{h}) = \arg \min_{q_2} \{\text{KL}(q_1 q_2|p)\} = \arg \min_{q_2} \{\mathcal{F}(q_1 q_2)\} \end{cases}$$

Summary of Bayesian methods for Blind Deconvolution

Alternate optimization for JMAP



Alternate optimization for VBA



JMAP and VBA with Gaussian priors

JMAP:

Initialization: $\mathbf{h}^{(0)} = \mathbf{h}_0$, $\mathbf{H} = \text{Convmtx}(\mathbf{h}^{(0)})$

Iterations:

$$\mathbf{f}^{(k)} = (\mathbf{H}'\mathbf{H} + \lambda_f \mathbf{I})^{-1} \mathbf{H}'\mathbf{g}$$

$$\mathbf{F} = \text{Convmtx}(\mathbf{f}^{(k-1)})$$

$$\mathbf{h}^{(k)} = (\mathbf{F}'\mathbf{F} + \lambda_h \mathbf{C}'_h \mathbf{C}_h)^{-1} \mathbf{F}'\mathbf{g}$$

$$\mathbf{H} = \text{Convmtx}(\mathbf{h}^{(k-1)})$$

VBA:

Initialization: $\mathbf{h}^{(0)} = \mathbf{h}_0$, $\mathbf{H} = \text{Convmtx}(\mathbf{h}^{(0)})$, $\mathbf{D}_f = \mathbf{I}$, $\mathbf{D}_h = \mathbf{I}$

Iterations:

$$\mathbf{f}^{(k)} = (\mathbf{H}'\mathbf{H} + \lambda_f \mathbf{I} + v_\epsilon \mathbf{D}'_f \mathbf{D}_f)^{-1} \mathbf{H}'\mathbf{g}$$

$$\Sigma_f = v_\epsilon (\mathbf{H}'\mathbf{H} + \lambda_f \mathbf{I} + v_\epsilon \mathbf{D}'_f \mathbf{D}_f)^{-1}$$

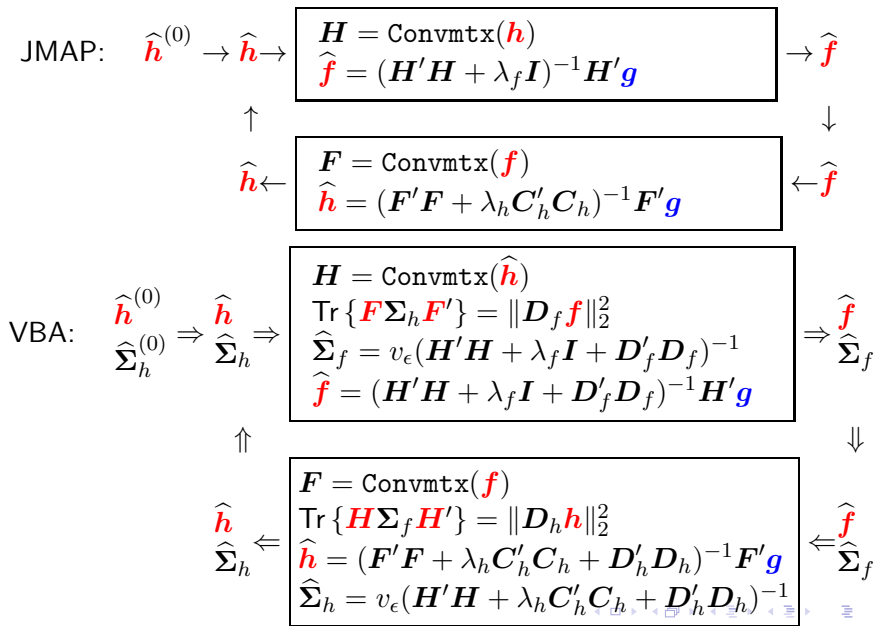
$$\mathbf{F} = \text{Convmtx}(\mathbf{f}^{(k-1)}) \quad \text{Tr}\{\mathbf{H}\Sigma_f\mathbf{H}'\} = \|\mathbf{D}'_f \mathbf{f}\|_2^2$$

$$\mathbf{h}^{(k)} = (\mathbf{F}'\mathbf{F} + \lambda_h \mathbf{C}'_h \mathbf{C}_h + v_\epsilon \mathbf{D}'_h \mathbf{D}_h)^{-1} \mathbf{F}'\mathbf{g}$$

$$\Sigma_h = v_\epsilon (\mathbf{F}'\mathbf{F} + \lambda_h \mathbf{C}'_h \mathbf{C}_h + \mathbf{D}'_h \mathbf{D}_h)^{-1}$$

$$\mathbf{H} = \text{Convmtx}(\mathbf{h}^{(k-1)}) \quad \text{Tr}\{\mathbf{F}\Sigma_h\mathbf{F}'\} = \|\mathbf{D}'_f \mathbf{f}\|_2^2$$

JMAP and VBA with Gaussian priors



Student-t prior

- ▶ Sparsity enforcing: Heavy tailed priors: Cauchy and Student-t:

$$p(\mathbf{f}) = \prod_j \mathcal{T}(f_j | \nu, v_f)$$

Infinite scaled mixture representation:

$$\mathcal{T}(f_j | \nu, v_f) = \int_0^\infty \mathcal{N}(f_j | 0, z_j^{-1}) \mathcal{G}(z_j | \nu/2, \nu/2) dz_j$$

- ▶ Hierarchical model: $p(\mathbf{f} | \mathbf{z}) p(\mathbf{z})$

$$p(\mathbf{f} | \mathbf{z}) = \prod_j \mathcal{N}(f_j | 0, z_j^{-1} v_f), \quad p(\mathbf{z}) = \prod_j \mathcal{G}(z_j | \alpha, \beta)$$

- ▶ Joint posterior: $p(\mathbf{h}, \mathbf{f}, \mathbf{z} | \mathbf{g}) \propto p(\mathbf{g} | \mathbf{h}, \mathbf{f}) p(\mathbf{f} | \mathbf{z}) p(\mathbf{z})$
 - ▶ JMAP: $(\hat{\mathbf{h}}, \hat{\mathbf{f}}, \hat{\mathbf{z}}) = \arg \max_{(\mathbf{h}, \mathbf{f}, \mathbf{z})} \{p(\mathbf{h}, \mathbf{f}, \mathbf{z} | \mathbf{g})\}$
 - ▶ VBA: Approximate $p(\mathbf{h}, \mathbf{f}, \mathbf{z} | \mathbf{g})$ by $q_1(\mathbf{h}) q_2(\mathbf{f}) q_3(\mathbf{z})$

JMAP and VBA with a Student-t prior

- ▶ Because we only changed $p(\mathbf{f})$ by $p(\mathbf{f}|\mathbf{z})p(\mathbf{z})$ where both are separable, the only changes are:

- ▶ replace

$$\mathbf{f}^{(k)} = (\mathbf{H}'\mathbf{H} + \lambda_f \mathbf{I})^{-1} \mathbf{H}'\mathbf{g}$$

- ▶ by:

$$\mathbf{f}^{(k)} = (\mathbf{H}'\mathbf{H} + \lambda_f \widehat{\mathbf{Z}}^{(k)})^{-1} \mathbf{H}'\mathbf{g}$$
$$\widehat{\mathbf{Z}}^{(k)} = \text{diag} \left[\widehat{z}_j^{(k)} \right], \quad \widehat{z}_j^{(k)} = \frac{\widehat{\beta}_j}{\widehat{\alpha}_j}$$

with

$$\begin{array}{ll} \text{JMAP} & \text{VBA} \\ \left\{ \begin{array}{l} \widehat{\alpha}_j = \alpha + \frac{1}{2} \|\mathbf{f}\|^2 \\ \text{and} \\ \widehat{\beta}_j = \beta + \frac{1}{2} + \|\mathbf{g} - \mathbf{h} * \mathbf{f}\|^2 \end{array} \right. & \left\{ \begin{array}{l} \widehat{\alpha}_j = \alpha + \frac{1}{2} \langle \|\mathbf{f}\|^2 \rangle \\ \text{and} \\ \widehat{\beta}_j = \beta + \frac{1}{2} + \langle \|\mathbf{g} - \mathbf{h} * \mathbf{f}\|^2 \rangle \end{array} \right. \end{array}$$

Comparison between JMAP and VBA

Alternate optimization for JMAP

$$\begin{array}{l}
 \hat{\mathbf{h}}^{(0)} \\
 \hat{\mathbf{z}}^{(0)}
 \end{array}
 \Rightarrow
 \begin{array}{l}
 \hat{\mathbf{h}} \\
 \hat{\mathbf{z}}
 \end{array}
 \Rightarrow
 \boxed{
 \begin{array}{l}
 \mathbf{Z} = \text{diag}[\hat{\mathbf{z}}], \quad \mathbf{H} = \text{Convmtx}(\hat{\mathbf{h}}) \\
 \hat{\mathbf{f}} = (\mathbf{H}'\mathbf{H} + \lambda_f \mathbf{Z})^{-1} \mathbf{H}'\mathbf{g}
 \end{array}
 }
 \rightarrow \hat{\mathbf{f}}$$

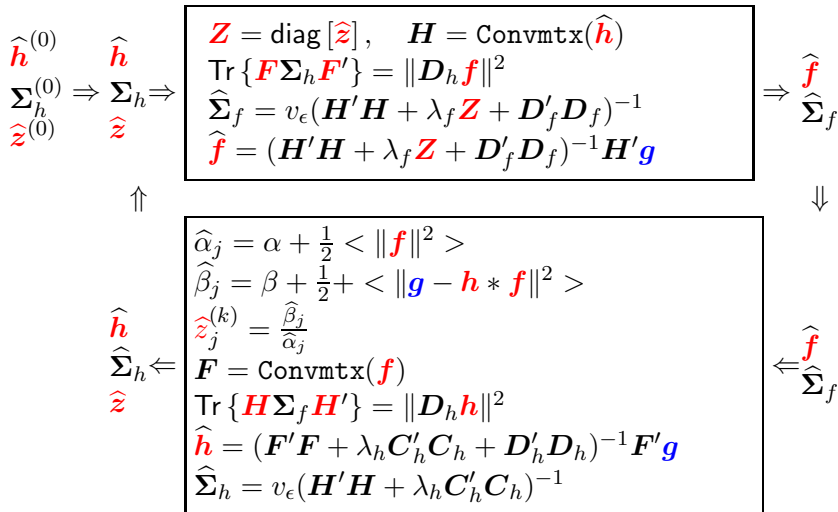
 \uparrow
 $\begin{array}{l} \hat{\mathbf{h}} \\ \hat{\mathbf{z}} \end{array} \leftarrow$

$$\boxed{
 \begin{array}{l}
 \hat{\alpha}_j = \alpha + \frac{1}{2} \|\mathbf{f}\|^2 \\
 \hat{\beta}_j = \beta + \frac{1}{2} + \|\mathbf{g} - \mathbf{h} * \mathbf{f}\|^2 \\
 \hat{\mathbf{z}}_j^{(k)} = \frac{\hat{\beta}_j}{\hat{\alpha}_j} \\
 \mathbf{F} = \text{Convmtx}(\mathbf{f}) \\
 \hat{\mathbf{h}} = (\mathbf{F}'\mathbf{F} + \lambda_h \mathbf{C}'_h \mathbf{C}_h)^{-1} \mathbf{F}'\mathbf{g}
 \end{array}
 }$$

 \downarrow
 $\leftarrow \hat{\mathbf{f}}$

Comparison between JMAP and VBA

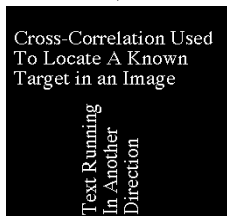
VBA



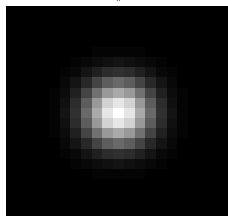
Conclusions

- ▶ In this paper, we considered the Blind Image Deconvolution problem in a Bayesian framework.
- ▶ We compared two main algorithms: JMAP and VBA giving some detailed insight for each of them for two cases:
 - ▶ Gaussian prior for both IRF h and the input signal f and
 - ▶ Gaussian prior for the IRF but a Student-t prior for the input signals or images to enhance or to account for possible sparsity structure of the input.
- ▶ JMAP: easy and not very costly but uncertainties are not accounted for.
- ▶ VBA: is more costly but both uncertainties of \hat{f} and \hat{h} are accounted for, in each iteration, for computing respectively, \hat{h} and \hat{f} .

Simulated results



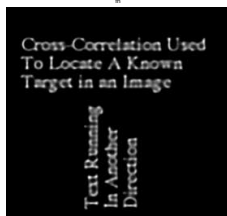
original f



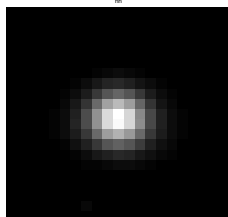
PSF h



Blurred & noisy g



deconv \hat{f}



Estimated PSF \hat{h}



Estimated \hat{f}

THANKS

Questions and Remarks