INSTRUMENT PARAMETER ESTIMATION IN BAYESIAN CONVEX DECONVOLUTION

F. Orieux [1], T. Rodet [1], J.-F. Giovannelli [2]

[1] Laboratoire des Signaux et Systèmes (CNRS – SUPELEC – Univ. Paris-Sud 11)
SUPELEC, Plateau de Moulon, 91 192 Gif-sur-Yvette Cedex, France, orieux@lss.supelec.fr.
[2] Laboratoire d'Intégration du Matériau au Système, Équipe Signal-Image,
Université de Bordeaux 1, 33405 Talence, France, Giova@IMS-Bordeaux.fr.

ABSTRACT

This paper proposes a Bayesian approach for estimation of instrument parameter in convex image deconvolution. The parameters of the instrument response (PSF) are jointly estimated with the image leading to a myopic deconvolution approach. In addition a special convex field allowing efficient hyperparameter estimation is used. The solution is based on a global *a posteriori* law for unknown parameters and object. The estimate is chosen in the sense of the posterior mean, numerically calculated by means of a Monte-Carlo Markov chain algorithm. The computation is efficient with a partial implementation in Fourier space. Simulation results are provided to assess the effectiveness of the proposed approach.

Index Terms— Instrument parameter estimation, myopic deconvolution, semi-blind deconvolution, convex deconvolution, Monte-Carlo Markov chain.

1. INTRODUCTION

Deconvolution constitutes an active research field as testified by the recent literature [1, 2]. Examples of application are medical imaging, astronomy, nondestructive testing and more generally imagery problems. In these applications, the degradation introduced by the instrument limits the interpretation of raw data whereas the need in resolution can be important.

Regarding processing methods, deconvolution raises an ill-posed problem and a solution relies on introduction of information in addition to the one provided by the data and the acquisition model. As a consequence, developed methods are specialized for the class of images in accordance with the introduced information. Form this stand point, this paper is dedicated to piecewise regular images with possible edges to be preserved or enhanced.

Anyway, the resulting estimate depends on variables in addition to the data. Firstly, the estimate naturally depends on the instrument response, namely the point spread function (PSF). The most part of the literature is devoted to the case of a known PSF, conversely, the present paper as well as more recent literature, is devoted to the case of poorly known PSF. In order to solve the problem, additional information must be introduced and two main configurations exist for the estimation.

The most common configuration operates when the shape of the PSF is known up to some parameters such width, orientation or position (shift or delay). The PSF equation is generally provided by the physical operating description. For example a Gaussian-shaped PSF is often used in optical imagery [3]. Parametric shape are also available in other acquisition modality: magnetic resonance force microscopy [4], spectrometry, optical microscopy [5],. . . The question of the estimation of this parameters jointly with the image is commonly called *myopic* deconvolution.

From the estimation point of view, the knowledge of the shape is a crucial information and is critical in order to resolve the ambiguity. As a counterpart, the observation model is generally in a non-linear dependency w.r.t to the instrument parameters. As a consequence, specific difficulties occur in the estimation process and one of the contribution of the proposed paper is to overcome these difficulties.

Secondly, the estimate also depends on the parameters of the probability laws named hyperparameters (mean, variance, parameters of correlation matrix,. . .) which tunes, among other things, the compromise between the information provided by the *a priori* and the information provided by the data. In real experiments, their values are unknown and the inversion problem including their estimation is referred to as *unsupervised*. The main difficulty is that the partition function of non-Gaussian Markov field are usually in unknown relation with the hyperparameters.The approach developed here is founded on a special correlated field based on a log-erf convex potential [6] and with an explicit partition function.

For the unsupervised convex blind deconvolution, Molina [1] tackle the question of the completly blind deconvolution problem and compute an approximation of the posterior law with a bayesian variationnal approach. For the myopic or semi-blind problem, Jalobeanu *et al.* [3] address the case of a symmetric Gaussian PSF. The width parameter and the noise variance are estimated in a preliminary step by Maximum-Likelihood based on the marginalization of the joint law. Since the probabilistic model are Gaussian, the integration is explicit and the maximisation is feasible.

In the present paper we address the convex myopic and unsupervised deconvolution problem. We propose a new method that jointly estimates the instrument parameter, the hyperparameter and the image with an extended *a posteriori* law for all unknown variables. The estimate is chosen as the mean of the posterior law, *i.e.,* the MMSE estimate and it is computed thanks to Monte-Carlo Markov Chain (MCMC) [7] stochastic sampling algorithm.

The paper is presented in the following manner. The two following sections describe our methodology: firstly the Bayesian probabilistic models are detailed in Sec. 2 and secondly, the MCMC algorithm used to compute the estimate, is described in Sec. 3. Numerical results are shown in Sec. 4. Finally, Sec. 5 is devoted to the conclusions and the perspectives.

2. BAYESIAN PROBABILISTIC MODEL

2.1. Direct model

We consider N pixels real square images represented in lexicographic order by vector $\boldsymbol{x} \in \mathbb{R}^N$, with generic elements x_n. The forward model is written $\boldsymbol{y} = \boldsymbol{H_w x} + \boldsymbol{n}$ where $\boldsymbol{y} \in \mathbb{R}^N$ is the data, $\boldsymbol{H_w}$ a convolution matrix parametrized by $\boldsymbol{w}$, $\boldsymbol{x}$ the image of interest and $\boldsymbol{n}$ the model errors. The parameters, such as width or orientation, of the parametric model (*e.g.,* Gaussian) are collected in vector $\boldsymbol{w} \in \mathbb{R}^P$. The matrix $\boldsymbol{H_w}$ is block-circulant circulant-block (BCCB) for computational efficiency of the convolution in Fourier space. The diagonalization of $\boldsymbol{H_w}$ is written $\boldsymbol{\Lambda_H} = \boldsymbol{F H_w F}^\dagger$ where $\boldsymbol{\Lambda_H}$ is a diagonal matrix, $\boldsymbol{F}$ the unitary Fourier matrix and $\dagger$ the transpose conjugate symbol.

2.2. Image prior law

In a general Gibbs form, the prior model writes

$$p\left(\boldsymbol{x}|\boldsymbol{\gamma}\right) = K_x(\boldsymbol{\gamma})^{-1}\ \exp\left[-\Phi_{\boldsymbol{\gamma}}\left(\boldsymbol{x}\right)\right]$$

where $\Phi_{\boldsymbol{\gamma}}$ is the energy controlled by a set of parameters (such as variance, threshold, correlation length…) collected in a vector $\boldsymbol{\gamma}$ and $K_x(\boldsymbol{\gamma})$ is the partition function. Usual approaches rely on Markov models: the energy involves local interactions $\Phi_{\boldsymbol{\gamma}}\left(\boldsymbol{x}\right) = \sum_p \varphi_{\boldsymbol{\gamma}}(\overline{x}_p)$ where $\overline{x}_p$ is a combination of pixels in the vicinity of pixel p. The complete design of the field relies on the structure of $\overline{\boldsymbol{x}}$ and the structure of φ.

As long as the question of parameter estimation is concerned, the key-point is the partition function K_x as a function of $\boldsymbol{\gamma}$, impossible for most models. Nevertheless, in a previous work [6], we have introduced a new model: it is the unique non-Gaussian Markov edge preserving field with an explicit partition function. It involves two variables: a pixel variable $\boldsymbol{x}$ and an auxiliary (or dual or hidden) variable $\boldsymbol{b}$ that catch the local spatial structure of $\boldsymbol{x}$. The joint law for $(\boldsymbol{x}, \boldsymbol{b})$ writes

$$p(\boldsymbol{x}, \boldsymbol{b}) = K_{x,b}^{-1} \exp\Big[-\gamma_{\boldsymbol{x}}\ \left\|\boldsymbol{Dx} - \boldsymbol{b}\right\|^2/2 + \gamma_{\boldsymbol{b}}\ \left\|\boldsymbol{b}\right\|_1/2\Big]$$

where $\|\boldsymbol{b}\|_1$ is the L_1 norm of $\boldsymbol{b}$, and the partition function

$$K_{x,b}^{-1} \propto (32\pi)^{-N/2}\ \gamma_{\boldsymbol{x}}^{N/2}\ \gamma_{\boldsymbol{b}}^{N}\,. \tag{1}$$

It presents two main advantages: 1) The partition function is explicit and easy to manipulate: it makes possible the development of efficient statistical methods for parameter estimation and 2) The conditional fields $p(\boldsymbol{x}|\boldsymbol{b})$ and $p(\boldsymbol{b}|\boldsymbol{x})$ are easy to simulate: the former is a (correlated) Gaussian component and the latter is a separable (non-Gaussian) component.

2.3. Noise and data laws

The statistical knowledge about the errors is modeled as white zero-mean Gaussian with unknown precision parameter γ_{n}. Consequently the parameter likelihood of the parameters attached to the data writes

$$p(\boldsymbol{y}|\boldsymbol{x}, \gamma_{\mathrm{n}}, \boldsymbol{w}) = (2\pi)^{-N/2}\gamma_{\mathrm{n}}^{N/2} \exp\left[-\frac{\gamma_{\mathrm{n}}}{2}\left\|\boldsymbol{y} - \boldsymbol{H_w x}\right\|^2\right].$$

It depends, of course, on the image $\boldsymbol{x}$, on the noise parameter γ_{n} and instrument parameters $\boldsymbol{w}$ embedded in $\boldsymbol{H_w}$. It naturally involves the least square term in spatial and spectral domain $\|\boldsymbol{y} - \boldsymbol{H_w x}\|^2 = \|\mathring{\boldsymbol{y}} - \boldsymbol{\Lambda_H}\mathring{\boldsymbol{x}}\|^2$. Clearly, the observation model $\boldsymbol{H_w}$ have generally a non-linear dependency w.r.t the parameters $\boldsymbol{w}$ that is transmitted to the estimator.

2.4. Hyperparameters law

The parameters $[\gamma_{\boldsymbol{x}}, \gamma_{\boldsymbol{b}}, \gamma_{\mathrm{n}}] = \boldsymbol{\gamma}$ are precisions of exponential family laws (Gaussian and Laplace laws). A conjugate law for this parameters is the Gamma law parametrized by two values (α_i, β_i) with $i = \boldsymbol{x}$, b or n

$$p(\gamma_i) = \frac{1}{\beta_i^{\alpha_i}\Gamma(\alpha_i)}\gamma^{\alpha_i - 1}\exp\left(-\gamma_i/\beta_i\right). \tag{2}$$

2.5. Instrument parameters law

Since the likelihood of the instrument parameter is intricate, no prior law allow easier calculation. Even if the noise law is Gaussian, the likelihood of the instrument parameter is not Gaussian at all (expect for linear relation like a gain).

In addition, generally nominal value with uncertainty comes with a parametrized shape of the observation model. We consider here that this information can be modelized with a parameter value in interval $[\boldsymbol{w}_m\ \boldsymbol{w}_M]$. The "Principle of Insufficient Reason" leads to a uniform prior on this interval

$$p(\boldsymbol{w}) = \mathcal{U}_{[\boldsymbol{w}_m\ \boldsymbol{w}_M]}(\boldsymbol{w}). \tag{3}$$

However, within the proposed framework, the choice is not limited and other laws, such as Gaussian, are possible, without fundamental changes in our methodology. We point out that whatever the prior law, manipulation and calculations remains difficult due to the non-linear dependency between the data and the instrument parameter. Consequently the choice of the prior law depends mainly on information modelization principles.

3. POSTERIOR MEAN ESTIMATOR AND LAW EXPLORATION

This section presents the algorithm to explore the posterior law and to compute an estimate of the parameters. For this purpose a Gibbs sampler [7] is used to provide samples.

3.1. Sampling the image

The conditional posterior law of the image (especially given the auxiliary variables) is Gaussian in Fourier space with inverse covariance matrix

$$\boldsymbol{\Sigma}^{(k+1)} = \gamma_{\mathrm{n}}^{(k)} |\boldsymbol{\Lambda}_{\boldsymbol{H}}^{(k)}|^2 + \gamma_{\boldsymbol{x}}^{(k)} |\boldsymbol{\Lambda}_{\boldsymbol{D}}|^2 \tag{4}$$

and mean

$$\boldsymbol{\mu}^{(k+1)} = \left(\boldsymbol{\Sigma}^{(k+1)}\right)^{-1} \left(\gamma_{\mathrm{n}}^{(k)} \boldsymbol{\Lambda}_{\boldsymbol{H}}^{*\,(k)} \mathring{\boldsymbol{y}} + \gamma_{\boldsymbol{x}}^{(k)} \boldsymbol{\Lambda}_{\boldsymbol{D}}^{*} \mathring{\boldsymbol{b}}\right). \tag{5}$$

The vector $\boldsymbol{\mu}^{(k+1)}$ is the regularized least square solution at current iteration (or the Wiener-Hunt solution).

3.2. Sampling auxiliary variables

The sampling of auxiliary variables is delicate but they are conditionally independent in spatial space. Consequently the sampling can be done very efficiently in parallel. In addition this special field allow sampling with the inversion of the cumulative density function (cdf) $F_{B|X}$ [6]. Because of numerical error, we use instead an efficient independent Metropolis-Hastings.

3.3. Sampling precision parameters

The conditional posterior laws of the precisions are Gamma corresponding to their prior law with parameters updated by the likelihood $\gamma_i^{(k+1)} \sim \mathcal{G}(\gamma_i|\alpha_i, \beta_i)$. In the case of Jeffreys's prior, the parameters are

$$\alpha_{\mathrm{n}} = N/2 \quad \text{and} \quad 2\beta_{\mathrm{n}}^{-1} = \left\|\mathring{\boldsymbol{y}} - \boldsymbol{\Lambda}_{\boldsymbol{H}}^{(k)} \mathring{\boldsymbol{x}}^{(k+1)}\right\|^2,$$

$$\alpha_{\mathrm{b}} = N \quad \text{and} \quad 2\beta_{\mathrm{b}}^{-1} = \left\|\boldsymbol{b}^{(k+1)}\right\|_1,$$

$$\alpha_{\boldsymbol{x}} = (N-1)/2 \quad \text{and} \quad 2\beta_{\boldsymbol{x}}^{-1} = \left\|\boldsymbol{\Lambda}_{\boldsymbol{D}} \mathring{\boldsymbol{x}}^{(k+1)}\right\|^2.$$

3.4. Sample instrument parameters

The conditional law for instrument parameters writes

$$\boldsymbol{w}^{(k+1)} \propto \exp\left[-\frac{\gamma_{\mathrm{n}}}{2}\left\|\mathring{\boldsymbol{y}} - \boldsymbol{\Lambda}_{\boldsymbol{H},\boldsymbol{w}}\,\mathring{\boldsymbol{x}}\right\|^2\right] \mathcal{U}_{[\boldsymbol{w}_m\ \boldsymbol{w}_M]}(\boldsymbol{w})$$

where parameters $\boldsymbol{w}$ are embedded in the instrument response $\boldsymbol{\Lambda}_{\boldsymbol{H}}$. This law is intricate and not standard, so no algorithm exists for direct sampling. Moreover, since this law mainly depend on $\boldsymbol{\Lambda}_{\boldsymbol{H}}$, or the parametric shape defined by the application, the choice of the prior law doesn't influence the complexity.

The Metropolis-Hastings (M.-H.) method is used to bypass this difficulty. In M.-H. algorithm, a sample $\boldsymbol{w}_{\mathrm{p}}$ is proposed and rejected or accepted with a prescribed probability. This probability depends on the ratio between the likelihood of the proposed value $\boldsymbol{w}_{\mathrm{p}}$ and the likelihood of the current value $\boldsymbol{w}^{(k)}$. Consequently the knowledge of the target law is only needed up to a normalisation constant. In addition it is only necessary to evaluate it at current state $\boldsymbol{w}_{\mathrm{p}}$ or $\boldsymbol{w}^{(k)}$ to use this algorithm.

3.5. Empirical mean

The sampling of $\mathring{\boldsymbol{x}}$, $\boldsymbol{b}$, $\boldsymbol{\gamma}$ and $\boldsymbol{w}$ are repeated iteratively until the law has been sufficiently explored. These samples $\left[\mathring{\boldsymbol{x}}^{(k)}, \boldsymbol{b}^{(k)}, \boldsymbol{\gamma}^{(k)}, \boldsymbol{w}^{(k)}\right]$ follow the global *a posteriori* law. By the large numbers law, the estimate, defined as the posterior mean, is approximated with $\hat{\boldsymbol{x}} = \boldsymbol{F}^{\dagger}\mathbb{E}[\mathring{\boldsymbol{x}}]$.

4. EXPERIMENTAL RESULTS

This section presents results obtained with the proposed approach on a realistic case Fig. 1(a).The method is also compared to [1], named BD, whose software is available on-line.

The instrument response $\boldsymbol{\Lambda}_{\boldsymbol{H}}$ is a discretized normalized Gaussian PSF written in Fourier space

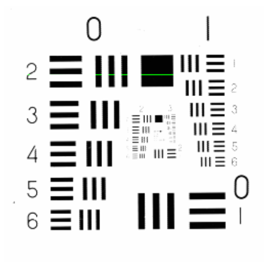
$$\mathring{h}(\nu_\alpha, \nu_\beta) = \exp\Bigg(-2\pi^2\Big(\nu_\alpha^2(w_\alpha \cos^2\varphi + w_\beta \sin^2\varphi) + \nu_\beta^2(w_\alpha \sin^2\varphi + w_\beta \cos^2\varphi) + 2\nu_\alpha\nu_\beta \sin\varphi\cos\varphi\,(w_\alpha - w_\beta)\Big)\Bigg)$$

with frequencies $(\nu_\alpha, \nu_\beta) \in [-0.5; 0.5]^2$. This low-pass filter is controlled by three parameters:

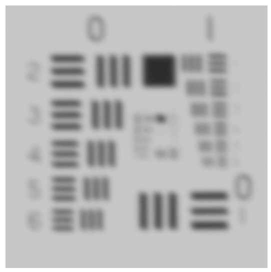
- two width parameters w_α and w_β. Their prior law interval are generally chosen to have an uncertainty of approximately ± 20% and 10% around a nominal value (see Sec 2.5).
- a rotation parameter φ set to $\pi/3$ with an *a priori* law $p(\varphi) = \mathcal{U}_{[\pi/4\ \pi/2]}(\varphi)$. The corresponding uncertainty is approximately ± 50% around the nominal value.

The images are square with 256×256 pixels. The matrix $\mathbf{\Lambda}_{\boldsymbol{D}}$ is obtained with the FFT-2D of the Laplacian.

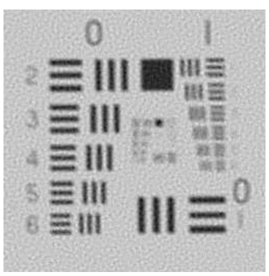
The Fig. 1(c) shows the obtained results with our approach for this image. Clearly the image is restored with more spatial details and more high spatial frequency than data Fig. 1(b). The Fig. 1(d) illustrates the result obtained with the BD approach, with limitation of 220 iterations. This image 1(d) suffer from the lack of knowledge of the PSF and the method tend to estimate a larger PSF up to the limited support.



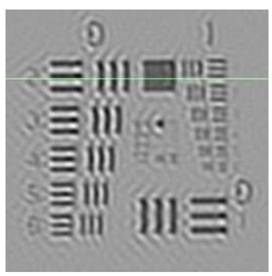
(a) True



(b) Data



(c) Proposed



(d) BD

Fig. 1. Fig. (a) is the true image of the Caterpillar Inc. test pattern and (b) the data. Fig. (c) is the estimate with the proposed approach. Fig. (d) correspond to the BD approach that estimate all the pixel of a non-parametric PSF on a limited support.

About the instrument parameters the estimated value are less close to the true value : $\widehat{w}_\alpha = 9.11$ for $w_\alpha = 11$ and $\widehat{w}_\beta = 5.9$ for $w_\beta = 7$. A possible interpretation is : since there is an anisotropy in the data because of the asymmetric PSF, the approach tends to explain the phenomena with anisotropy both in the image and the PSF.

	True (prior interval)	Our approach	BD
w_α	11 ([8 12])	9.11 ± 0.0083	-
w_β	7 ([5 8])	5.9 ± 0.01	-
φ	1.047	1.157 ± 0.001	-
$error$	-	0.01 %	43.27 %

Table 1. Instruments parameters estimation. Estimation of the parameter is not available in the non-parametric BD approach. The error is on the energy of the PSF.

About the hyper-parameter estimation since the true values are not known it is not possible to make a comparison, expect for the noise value. For this parameter, the estimated value is very close to the true value with $\widehat{\gamma}_{\mathrm{n}} = 1.0005 \times 10^6$ for $\gamma_{\mathrm{n}} = 10^6$. The interpretation is that a lot of information is available in data for the noise since it is in the output of the system.

5. CONCLUSION

This paper presents a new global and coherent method for unsupervised myopic robust deconvolution. It is build within a Bayesian framework and an extended *a posteriori* law for the instrument parameters, the image and the hyperparameters. The estimate, defined as the posterior mean, is computed by means of an MCMC algorithm. A parametric instrument response and an automatic balance between data information and *a priori* information are jointly estimated with the image. In addition the results show that the deconvolved image is closer to the true image than the data and show restored high-frequencies.

6. REFERENCES

[1] R. Molina, J. Mateos, and A. K. Katsaggelos, "Blind deconvolution using a variational approach to parameter, image, and blur estimation," *IEEE Trans. Image Processing*, vol. 15, no. 12, pp. 3715–3727, December 2006.

[2] P. Campisi and K. Egiazarian, Eds., *Blind Image Deconvolution*, CRC Press, 2007.

[3] A. Jalobeanu, L. Blanc-Feraud, and J. Zerubia, "Estimation of blur and noise parameters in remote sensing," in *Proc. IEEE ICASSP*, 2002, vol. 4, pp. 3580–3583.

[4] N. Dobigeon, A.O. Hero, and J.-Y. Tourneret, "Hierarchical bayesian sparse image reconstruction with application to MRFM," *IEEE Trans. Image Processing*, September 2009.

[5] Praveen Pankajakshani, Bo Zhang, Laure Blanc-Féraud, Zvi Kam, Jean-Christophe Olivo-Marin, and Josiane Zerubia, "Blind deconvolution for thin-layered confocal imaging," *Appl. Opt.*, vol. 48, no. 22, pp. 4437–4448, 2009.

[6] Jean-François Giovannelli, "Unsupervised bayesian convex deconvolution based on a field with an explicit partition function," *IEEE Trans. Image Processing*, vol. 17, no. 1, pp. 16–26, January 2008.

[7] Christian P. Robert and George Casella, *Monte-Carlo Statistical Methods*, Springer Texts in Statistics. Springer, New York, NY, 2000.