Hyperparameter estimation in image restoration problems with partially known blurs

Nikolas P. Galatsanos
Illinois Institute of Technology
Department of Electrical and Computer Engineering
3301 South Dearborn Street
Chicago, Illinois 60616

Vladimir Z. Mesarović
Crystal Semiconductor Corporation
4210 South Industrial Drive
Austin, Texas 78744

Rafael Molina, MEMBER SPIE
Universidad de Granada
Departamento de Ciencias de la Computación e I. A.
E-18071 Granada, Spain
E-mail: rms@decsai.ugr.es

Aggelos K. Katsaggelos
Northwestern University
Department of Electrical and Computer Engineering
Evanston, Illinois 60208-3118

Javier Mateos
Universidad de Granada
Departamento de Ciencias de la Computación e I. A.
E-18071 Granada, Spain

Abstract. This work is motivated by the observation that it is not possible to reliably estimate simultaneously all the necessary hyperparameters in an image restoration problem when the point-spread function is assumed to be the sum of a known deterministic and an unknown random component. To solve this problem we propose to use gamma hyperpriors for the unknown hyperparameters. Two iterative algorithms that simultaneously restore the image and estimate the hyperparameters are derived, based on the application of evidence analysis within the hierarchical Bayesian framework. Numerical experiments are presented that show the benefits of introducing hyperpriors for this problem. [DOI: 10.1117/1.1487850]

Subject terms: image restoration; partially known blur; hyperparameter estimation; gamma hyperpriors; Bayesian estimation.

Paper 010263 received July 31, 2001; revised manuscript received Jan. 29, 2002; accepted for publication Feb. 4, 2002.

1 Introduction

Traditionally, image restoration algorithms have assumed exact knowledge of the blurring operator. In recent years, in particular in the field of astronomical image restoration,[1,2] a significant effort has been devoted to solving the so-called blind deconvolution problem, in which it is assumed that little or nothing is known about the underlying blurring process (see Refs. 3,4 and references therein). In most practical applications, the point-spread function (PSF) is neither unknown nor perfectly known,[5] that is, usually some information about the PSF is available, but this information is never exact.

The use of a PSF modeled by a known mean and an additive random error component has been addressed in the past (see, for instance, Refs. 6–8). However, in all these works the needed model parameters were assumed known. More recently, attempts were made to address the parameter estimation problem: in Refs. 9,10 and 11 (Chapter III) the expectation-maximization algorithm was used, and in Refs. 11 (Chapter IV) and 12–14 the estimation was addressed within the hierarchical Bayesian[15] framework. However, in Refs. 9,10, and 11 (Chapters III and IV), and 12–14 we observed that it was not possible to reliably estimate *simultaneously* the hyperparameters that capture the variances of the PSF error and the additive noise.

In this paper we ameliorate the difficulties of estimating all the necessary hyperparameters by introducing gamma hyperpriors within the hierarchical Bayesian framework. We derive two iterative algorithms that simultaneously estimate all the necessary hyperparameters and restore the image.

The rest of this paper is organized as follows: In Sec. 2 the image model, two models for the fidelity to the data, and the hyperparameter model are discussed. In Sec. 3 the basic philosophy behind evidence analysis (EA) is briefly presented and its application to the restoration problem from partially known blur is discussed. Section IV presents two EA algorithms using the different proposed models. In Sec. 5 we present numerical experiments that compare the proposed approaches. Section 6 concludes the paper.

2 Components of the Hierarchical Model

Let us now examine the components of the hierarchical model used for the restoration problem with partially known blur, that is, the image model, the observation model, and the model for the unknown hyperparameters.

A commonly used model for the image prior in image restoration problems is the simultaneously autoregressive

(SAR) model.[16] This model can be described by the following conditional PDF:

$$P(\mathbf{f}|\alpha)=\text{const}\cdot\alpha^{N/2}\exp\left\{-\frac{\alpha}{2}\|\mathbf{Qf}\|^2\right\}, \tag{1}$$

where $\mathbf{f}\in\mathcal{R}^N$ represents the source image and α is an unknown positive parameter that controls the smoothness of the image. For simplicity, but without loss of generality, we shall use a circulant Laplacian high-pass operator for $\mathbf{Q}$ throughout the rest of this paper.

The space-invariant PSF is represented as the sum of a deterministic component and a stochastic component of zero mean, i.e.,

$$\mathbf{h}=\overline{\mathbf{h}}+\mathbf{\Delta h}, \tag{2}$$

where $\overline{\mathbf{h}}\in\mathcal{R}^N$ is the deterministic (known) component of the PSF and $\mathbf{\Delta h}\in\mathcal{R}^N$ is the random (unknown error) component modeled as zero-mean white noise with covariance matrix $\mathbf{R}_{\Delta h}=(1/\beta)\,\mathbf{I}_{N\times N}$. For our problem, the image degradation can be described, in lexicographical form, by the model[6–8]

$$\mathbf{g}=\mathbf{Hf}+\mathbf{\Delta g}, \tag{3}$$

in which

$$\mathbf{H}=\overline{\mathbf{H}}+\mathbf{\Delta H}, \tag{4}$$

where $\mathbf{g},\mathbf{\Delta g}\in\mathcal{R}^N$ represent, respectively, the observed degraded image and the additive zero-mean white noise in the observed image, with covariance matrix $\mathbf{R}_{\Delta g}=(1/\gamma)\,\mathbf{I}_{N\times N}$. The matrix $\overline{\mathbf{H}}$ is the known (assumed, estimated, or measured) component of the $N\times N$ PSF matrix $\mathbf{H}$; $\mathbf{\Delta H}$ is the unknown component of $\mathbf{H}$, generated by $\mathbf{\Delta h}$ defined in Eq. (2).

From Eqs. (2)–(4) it is clear that the form of the conditional distribution of $\mathbf{g}$ is not simple. In fact we are going to propose two different models for $P(\mathbf{g}|\mathbf{f},\alpha,\beta,\gamma)$.

For the *fixed*-**f** *covariance model* we assume that both the PSF noise $\mathbf{\Delta h}$ and the additive noise $\mathbf{\Delta g}$ are Gaussian. Then, since the vector $\mathbf{f}$ is not a random quantity but rather a fixed one, it is straightforward to see from Eq. (3) that $P(\mathbf{g}|\mathbf{f},\alpha,\beta,\gamma)$ is given by

$$P(\mathbf{g}|\mathbf{f},\alpha,\beta,\gamma)\propto[\det(\mathbf{R}_{g|f})]^{-1/2} \times\exp[-\tfrac{1}{2}(\mathbf{g}-\overline{\mathbf{H}}\mathbf{f})^t\mathbf{R}_{g|f}^{-1}(\mathbf{g}-\overline{\mathbf{H}}\mathbf{f})]. \tag{5}$$

The conditional covariance $\mathbf{R}_{g|f}$ in Eq. (5) is given by[11,17]

$$\mathbf{R}_{g|f}=\mathbf{F}\mathbf{R}_{\Delta h}\mathbf{F}^t+\mathbf{R}_{\Delta g}=\frac{1}{\beta}\mathbf{FF}^t+\frac{1}{\gamma}\mathbf{I}, \tag{6}$$

where we have used the commutative property of the convolution and $\mathbf{F}$ denotes the circulant matrix generated by the image $\mathbf{f}$; see Ref. 11,14 for details.

For the *averaged*-**f** *covariance model* we assume that the *observations* **g** are Gaussian, and instead of using $\mathbf{FF}^t$ in the expression for the covariance we use its mean value from the prior. Thus, for this model we get

$$P(\mathbf{g}|\mathbf{f},\alpha,\beta,\gamma)\propto[\det(\underline{\mathbf{R}}_{g|f})]^{-1/2} \times\exp\{-\tfrac{1}{2}(\mathbf{g}-\overline{\mathbf{H}}\mathbf{f})^t\underline{\mathbf{R}}_{g|f}^{-1}(\mathbf{g}-\overline{\mathbf{H}}\mathbf{f})\}. \tag{7}$$

where

$$\underline{\mathbf{R}}_{g|f}=\frac{N}{\beta}(\alpha\mathbf{Q}^t\mathbf{Q})^{-1}+\frac{1}{\gamma}\mathbf{I}. \tag{8}$$

Note that by using this approximation we have incorporated the uncertainty of the image prior model, α, in the conditional distribution, which made the $\log P(\mathbf{g}|\mathbf{f},\alpha,\beta,\gamma)$ function quadratic with respect to $\mathbf{f}$. This yields a linear estimator for $\mathbf{f}$, as will be shown in the following section.

The Bayesian formulation allows the introduction of information about parameters that have to be estimated by using prior distributions over them.[15] To do so we use, as hyperprior, the gamma distribution defined by

$$P(x)\propto x^{(l_x-2)/2}\exp\{-m_x(l_x-2)x\}, \tag{9}$$

where $x\in\{\alpha,\beta,\gamma\}$ denotes a hyperparameter, and the parameters l_x and m_x are explained below. The mean and the variance of a random variable x with PDF in Eq. (9) are given by

$$E\{x\}=\frac{l_x}{2m_x(l_x-2)}\qquad\text{and}\qquad\text{Var}\{x\}=\frac{l_x}{2m_x^2(l_x-2)^2}. \tag{10}$$

According to Eq. (10), for l_x large, the mean of x is approximately equal to $1/2m_x$, and its variance decreases when l_x increases. Thus, $1/2m_x$ specifies the mean of the gamma distributed random variable x, while l_x can be used as a measure of the certainty in the knowledge about this mean.

3 Hierarchical Bayesian Analysis

In this work, the joint distribution we use is defined by

$$P(\mathbf{g},\mathbf{f},\alpha,\beta,\gamma) =P(\mathbf{g}|\mathbf{f},\alpha,\beta,\gamma)P(\mathbf{f}|\alpha,\beta,\gamma)P(\alpha)P(\beta)P(\gamma). \tag{11}$$

To estimate the unknown hyperparameters and the original image, we apply evidence analysis, since we have found that it provides good results for restoration-reconstruction problems.[18] According to the EA approach the simultaneous estimation of **f**, α, β, and γ is performed as follows:

- Parameter estimation step:

$$\hat{\alpha},\hat{\beta},\hat{\gamma}=\arg\max_{\alpha,\beta,\gamma}\{P(\alpha,\beta,\gamma|\mathbf{g})\}$$
$$=\arg\max_{\alpha,\beta,\gamma}\left\{\int_{\mathbf{f}}P(\mathbf{g},\mathbf{f},\alpha,\beta,\gamma)\mathrm{d}\mathbf{f}\right\}. \tag{12}$$

- Restoration step:

$$\hat{\mathbf{f}}(\hat{\alpha},\hat{\beta},\hat{\gamma})=\arg\max_{\mathbf{f}}\{P(\mathbf{f}|\mathbf{g},\hat{\alpha},\hat{\beta},\hat{\gamma})\}$$
$$=\arg\max_{\mathbf{f}}\{P(\mathbf{f}|\hat{\alpha},\hat{\beta},\hat{\gamma})P(\mathbf{g}|\mathbf{f},\hat{\alpha},\hat{\beta},\hat{\gamma})\}, \tag{13}$$

The estimates $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\gamma}$ from the parameter estimation step depend on the current estimate of the image. Likewise, the estimate $\hat{\mathbf{f}}$ from the restoration step will depend on the current estimates of the parameters. Therefore, the above two-step procedure is repeated until convergence occurs.

Using the two different choices for $P(\mathbf{g}|\mathbf{f},\alpha,\beta,\gamma)$ given in Eqs. (5) and (7), we will now proceed with the evidence analysis.

4 Proposed Algorithms

4.1 *Evidence Analysis Based on the Fixed-f Covariance Model*

Substituting Eqs. (1) and (5) into (11), we obtain

$$P(\mathbf{g},\mathbf{f},\alpha,\beta,\gamma)\approx\alpha^{N/2}[\det(\mathbf{R}_{g|f})]^{-1/2}$$
$$\times\exp[-\tfrac{1}{2}J(\mathbf{f},\alpha,\beta,\gamma)]P(\alpha)P(\beta)P(\gamma), \tag{14}$$

where

$$J(\mathbf{f},\alpha,\beta,\gamma)=\alpha\|\mathbf{Q}\mathbf{f}\|^2+(\mathbf{g}-\overline{\mathbf{H}}\mathbf{f})^t\mathbf{R}_{g|f}^{-1}(\mathbf{g}-\overline{\mathbf{H}}\mathbf{f}). \tag{15}$$

4.1.1 *Parameter estimation step*

To estimate $\hat{\alpha},\hat{\beta},\hat{\gamma}$, we first have to integrate $P(\mathbf{g},\mathbf{f},\alpha,\beta,\gamma)$ in Eq. (14) over $\mathbf{f}$, that is,

$$P(\alpha,\beta,\gamma|\mathbf{g})\propto P(\alpha)P(\beta)P(\gamma)\alpha^{N/2}\int_{\mathbf{f}}[\det(\mathbf{R}_{g|f})]^{-1/2}$$
$$\times\exp\{-\tfrac{1}{2}J(\mathbf{f},\alpha,\beta,\gamma)\}d\mathbf{f}. \tag{16}$$

To perform the integration in Eq. (16), we expand $J(\mathbf{f},\alpha,\beta,\gamma)$ in Taylor series around a known $\mathbf{f}^{(n)}$, where (n) denotes the iteration index, i.e.,

$$J(\mathbf{f},\alpha,\beta,\gamma)\approx J(\mathbf{f}^{(n)},\alpha,\beta,\gamma)+(\mathbf{f}-\mathbf{f}^{(n)})^t\nabla J(\mathbf{f},\alpha,\beta,\gamma)|_{\mathbf{f}^{(n)}}$$
$$+\tfrac{1}{2}(\mathbf{f}-\mathbf{f}^{(n)})^t\nabla^2 J(\mathbf{f},\alpha,\beta,\gamma)|_{\mathbf{f}^{(n)}}(\mathbf{f}-\mathbf{f}^{(n)}). \tag{17}$$

Note that, in Eq. (17),

$$\nabla J(\mathbf{f},\alpha,\beta,\gamma)|_{\mathbf{f}^{(n)}}=0 \tag{18}$$

if $\mathbf{f}^{(n)}$ is chosen to be the minimizer of $J(\mathbf{f},\alpha,\beta,\gamma)$ in Eq. (15), and that the Hessian matrix can be approximated by

$$\nabla^2 J(\mathbf{f},\alpha,\beta,\gamma)|_{\mathbf{f}^{(n)}}=2\mathbf{G}^{(n)}=2(\alpha\mathbf{Q}^t\mathbf{Q}+\overline{\mathbf{H}}^t\mathbf{R}_{g|f^{(n)}}^{-1}\overline{\mathbf{H}}), \tag{19}$$

where we have not taken into account the derivatives of $\mathbf{R}_{g|f^{(n)}}^{-1}$ with respect to $\mathbf{f}$.[11,14,17]

Finally, substituting Eq. (17) into (16), and using the fact that the factor $[\det(\mathbf{R}_{g|f})]^{-1/2}$ can be replaced by $[\det(\mathbf{R}_{g|f^{(n)}})]^{-1/2}$ because it depends weakly on $\mathbf{f}$ compared to the exponential term under the integral,[11,17] Eq. (16) becomes

$$P(\alpha,\beta,\gamma|\mathbf{g})\propto P(\alpha)P(\beta)P(\gamma)\alpha^{N/2}\det[\mathbf{R}_{g|f^{(n)}}]^{-1/2}$$
$$\times\det[\mathbf{G}^{(n)}]^{-1/2}\exp[-\tfrac{1}{2}J(\mathbf{f}^{(n)},\alpha,\beta,\gamma)]. \tag{20}$$

Taking 2 log on both sides of Eq. (20), and differentiating with respect to α,β,γ, we obtain the following iterations:

$$\frac{1}{\alpha^{(n+1)}}=2m_\alpha\mu_\alpha+(1-\mu_\alpha)\frac{1}{N}[\|\mathbf{Q}\mathbf{f}^{(n)}\|^2$$
$$+\mathrm{tr}(\mathbf{G}^{(n)\,-1}\mathbf{Q}^t\mathbf{Q})], \tag{21}$$

$$\frac{1}{\beta^{(n+1)}}=2m_\beta\mu_\beta+(1-\mu_\beta)\frac{1}{N}$$
$$\times\left[(\mathbf{g}-\overline{\mathbf{H}}\mathbf{f}^{(n)})^t\mathbf{R}_{g|f^{(n)}}^{-1}\frac{1}{\beta^{(n)2}}\mathbf{F}^{(n)}\mathbf{F}^{(n)\,t}\right.$$
$$\times\mathbf{R}_{g|f^{(n)}}^{-1}(\mathbf{g}-\overline{\mathbf{H}}\mathbf{f}^{(n)})+\mathrm{tr}\left(\frac{1}{\beta^{(n)}\gamma^{(n)}}\mathbf{R}_{g|f^{(n)}}^{-1}\right)$$
$$\left.+\mathrm{tr}\left(\mathbf{G}^{(n)\,-1}\overline{\mathbf{H}}^t\mathbf{R}_{g|f^{(n)}}^{-1}\frac{1}{\beta^{(n)2}}\mathbf{F}^{(n)}\mathbf{F}^{(n)t}\mathbf{R}_{g|f^{(n)}}^{-1}\overline{\mathbf{H}}\right)\right], \tag{22}$$

$$\frac{1}{\gamma^{(n+1)}}\frac{1}{N}=2m_\gamma\mu_\gamma+(1-\mu_\gamma)\left[(\mathbf{g}-\overline{\mathbf{H}}\mathbf{f}^{(n)})^t\frac{1}{\gamma^{(n)2}}\mathbf{R}_{g|f^{(n)}}^{-2}\right.$$
$$\times(\mathbf{g}-\overline{\mathbf{H}}\mathbf{f}^{(n)})+\mathrm{tr}\left(\frac{1}{\beta^{(n)}\gamma^{(n)}}\mathbf{F}^{(n)}\mathbf{F}^{(n)\,t}\mathbf{R}_{g|f^{(n)}}^{-1}\right)$$
$$\left.+\mathrm{tr}\left(\mathbf{G}^{(n)\,-1}\overline{\mathbf{H}}^t\frac{1}{\gamma^{(n)2}}\mathbf{R}_{g|f^{(n)}}^{-2}\overline{\mathbf{H}}\right)\right]. \tag{23}$$

where the normalized confidence parameter μ_x, $x\in\{\alpha,\beta,\gamma\}$, is defined as

$$\mu_x=1-\frac{N}{N+l_x-2}. \tag{24}$$

4.1.2 *Restoration step*

According to Eq. (13),

$$\hat{\mathbf{f}}(\hat{\alpha},\hat{\beta},\hat{\gamma})=\arg\max_{\mathbf{f}} P(\mathbf{f}|\mathbf{g},\hat{\alpha},\hat{\beta},\hat{\gamma})$$
$$=\arg\min_{\mathbf{f}}\{(\overline{\mathbf{H}}\mathbf{f}-\mathbf{g})^t\hat{\mathbf{R}}_{g|f}^{-1}(\overline{\mathbf{H}}\mathbf{f}-\mathbf{g})+\hat{\alpha}\|\mathbf{Q}\mathbf{f}\|^2$$
$$+\log[\det(\hat{\mathbf{R}}_{g|f})]\}, \tag{25}$$

where $\hat{\mathbf{R}}_{g|f}=(1/\hat{\beta})\,\mathbf{F}\mathbf{F}^t+(1/\hat{\gamma})\,\mathbf{I}$. The functional in Eq. (25) is nonconvex and may have several local minima. In general, a closed-form solution to Eq. (25) does not exist and numerical optimization algorithms must be used. A practical computation of Eq. (25) can be obtained by transforming it to the DFT domain.[11,14]

4.2 *Evidence Analysis Based on the Averaged-***f** *Covariance Model*

Using Eq. (7) as the likelihood equation, we derive another iterative parameter-estimation–image-restoration algorithm for this problem. We follow identical steps as in the previous section with

$$\underline{\mathbf{R}}_{g|f}=\frac{N}{\beta}(\alpha\mathbf{Q}^t\mathbf{Q})^{-1}+\frac{1}{\gamma}\mathbf{I}. \tag{26}$$

4.2.1 *Parameter estimation step*

To find the estimates of the parameters, $P(\alpha,\beta,\gamma|\mathbf{g})$ must be maximized. Taking the derivatives with respect to α, β, γ gives the following iterations:

$$\frac{1}{\alpha^{(n+1)}}=2m_\alpha\mu_\alpha+(1-\mu_\alpha)\Bigg(\frac{1}{N}\|\mathbf{Q}\mathbf{f}^{(n)}\|^2$$
$$-\mathrm{tr}\left[\underline{\mathbf{R}}_{g|f^{(n)}}^{-1}\frac{1}{\beta^{(n)}\alpha^{(n)\,2}}(\mathbf{Q}^t\mathbf{Q})^{-1}\right]$$
$$+\mathrm{tr}\Bigg[\underline{\mathbf{R}}_{g|f^{(n)}}^{-2}\frac{1}{\beta^{(n)}\alpha^{(n)\,2}}\,(\mathbf{Q}^t\mathbf{Q})^{-1}(\mathbf{g}-\overline{\mathbf{H}}\mathbf{f}^{(n)})$$
$$\times(\mathbf{g}-\overline{\mathbf{H}}\mathbf{f}^{(n)})^t\Bigg]+\mathrm{tr}\Bigg\{\underline{\mathbf{G}}^{(n)\;-1}$$
$$\times\left[\mathbf{Q}^t\mathbf{Q}+\overline{\mathbf{H}}^t\overline{\mathbf{H}}\underline{\mathbf{R}}_{g|f^{(n)}}^{-2}\frac{1}{\beta^{(n)}\alpha^{(n)\,2}}\,(\mathbf{Q}^t\mathbf{Q})^{-1}\right]\Bigg\}\Bigg), \tag{27}$$

$$\frac{1}{\beta^{(n+1)}}=2m_\beta\mu_\beta+(1-\mu_\beta)$$
$$\times\Bigg\{\mathrm{tr}\Bigg[\underline{\mathbf{R}}_{g|f^{(n)}}^{-2}\frac{1}{\beta^{(n)\,2}\,\alpha^{(n)}}(\mathbf{Q}^t\mathbf{Q})^{-1}(\mathbf{g}-\overline{\mathbf{H}}\mathbf{f}^{(n)})$$
$$\times(\mathbf{g}-\overline{\mathbf{H}}\mathbf{f}^{(n)})^t\Bigg]+\frac{1}{N\beta^{(n)}\gamma^{(n)}}\,\mathrm{tr}(\underline{\mathbf{R}}_{g|f^{(n)}}^{-1})$$
$$+\mathrm{tr}\left[\underline{\mathbf{G}}^{(n)^{-1}}\overline{\mathbf{H}}^t\overline{\mathbf{H}}\underline{\mathbf{R}}_{g|f^{(n)}}^{-2}\frac{1}{\beta^{(n)\,2}\alpha^{(n)}}(\mathbf{Q}^t\mathbf{Q})^{-1}\right]\Bigg\}, \tag{28}$$

$$\frac{1}{\gamma^{(n+1)}}=2m_\gamma\mu_\gamma+(1-\mu_\gamma)\,\frac{1}{N}$$
$$\times\Bigg\{\frac{1}{\beta^{(n)}\gamma^{(n)}}\,\mathrm{tr}[\underline{\mathbf{R}}_{g|f^{(n)}}^{-1}(\alpha^{(n)}\mathbf{Q}^t\mathbf{Q})^{-1}]$$
$$-\frac{1}{\gamma^{(n)\,2}}\,\mathrm{tr}(\underline{\mathbf{G}}^{(n)^{-1}}\overline{\mathbf{H}}^t\overline{\mathbf{H}}\underline{\mathbf{R}}_{g|f^{(n)}}^{-2})$$
$$-\frac{1}{\gamma^{(n)\,2}}\,\mathrm{tr}[\underline{\mathbf{R}}_{g|f^{(n)}}^{-2}(\mathbf{g}-\overline{\mathbf{H}}\mathbf{f}^{(n)})(\mathbf{g}-\overline{\mathbf{H}}\mathbf{f}^{(n)})^t]\Bigg\}, \tag{29}$$

where μ_x was defined in Eq. (24) and

$$\underline{\mathbf{G}}^{(n)}=\alpha^{(n)}\mathbf{Q}^t\mathbf{Q}+\overline{\mathbf{H}}^t\underline{\mathbf{R}}_{g|f^{(n)}}^{-1}\overline{\mathbf{H}}. \tag{30}$$

4.2.2 *Image restoration step*

For the image estimation step, similar to Eq. (25), we can write

$$\hat{\mathbf{f}}(\hat{\alpha},\hat{\beta},\hat{\gamma})=\arg\min_{\mathbf{f}}[(\overline{\mathbf{H}}\mathbf{f}-\mathbf{g})^t\hat{\underline{\mathbf{R}}}_{g|f}^{-1}(\overline{\mathbf{H}}\mathbf{f}-\mathbf{g})+\hat{\alpha}\|\mathbf{Q}\mathbf{f}\|^2]. \tag{31}$$

Note that, since $\hat{\underline{\mathbf{R}}}_{g|f}$ does not depend on $\mathbf{f}$, $P(\mathbf{g}|\mathbf{f},\hat{\alpha},\hat{\beta},\hat{\gamma})$ is quadratic with respect to $\mathbf{f}$. As a result, the image restoration step gives the linear estimate for $\mathbf{f}$

$$\hat{\mathbf{f}}(\hat{\alpha},\hat{\beta},\hat{\gamma})=(\overline{\mathbf{H}}^t\hat{\underline{\mathbf{R}}}_{g|f}^{-1}\overline{\mathbf{H}}+\hat{\alpha}\mathbf{Q}^t\mathbf{Q})^{-1}\overline{\mathbf{H}}^t\hat{\underline{\mathbf{R}}}_{g|f}^{-1}\mathbf{g}. \tag{32}$$

4.3 *Comments on the Normalized Confidence Parameters*

There is a very interesting and intuitive interpretation of the use of the normalized confidence parameters μ_x, $x\in\{\alpha,\beta,\gamma\}$ [see Eq. (24)] in the hyperparameter estimation procedures in Secs. 4.1.1 and 4.2.1. We are estimating the unknown hyperparameters by linearly weighting their maximum likelihood (ML) estimates with the prior knowledge we have on their means and variances. So, if we know, let us say from previous experience, the noise or image variances with some degree of certainty, we can use this knowledge to guide the convergence of the iterative procedures.

Note that, for a given hyperparameter $x\in\{\alpha,\beta,\gamma\}$, $\mu_x=0$ means having no confidence on the mean of its hyperprior, so that the estimate of x is identical to the ML estimate in Ref. 14. In contrast, when $\mu_x=1$, $x\in\{\alpha,\beta,\gamma\}$, the value of x is fixed to the mean of its hyperprior.

As will become evident in the following section, an important observation is that neither of the two algorithms can reliably estimate the PSF and additive noise variances β^{-1} and γ^{-1} simultaneously when no prior knowledge is introduced, that is, when $\mu_\alpha=\mu_\beta=\mu_\gamma=0.0$. This was expected, since the sum of these noises appears in the data. However, introducing some (even very little) prior knowledge about one of the parameter aids both methods to accurately estimate both β and γ while restoring the image.

Notice that if we do not have any prior knowledge about the values of the hyperparameters, we can first assume that there is no random component in the blur, that is, β^{-1}

Table 1 Comparison between EA1 and EA2 algorithms.

Algorithm	Restoration step	Parameter estimation step	Computational complexity (s/iteration)
EA1	Requires iterative optimization	Complicated relation with f	1.24
EA2	Linear closed-form solution	Quadratic dependence on f	0.07

$=0$, and estimate γ and α by any classical method, such as MLE. Once an estimate of γ is obtained (which is in general a good one, since the noise model is usually accurate, while the prior image model is an approximation and the PSF variance is typically very small), it is used with a medium to high confidence parameter to guide the estimation of the partially known PSF with $\mu_\alpha = \mu_\beta = 0.0$. Alternatively, we can run the algorithms with no prior knowledge of the parameters, that is, $\mu_\alpha = \mu_\beta = \mu_\gamma = 0.0$, using then a high confidence for the obtained γ^{-1} value ($\mu_\gamma \geqslant 0.7$) and letting the algorithms to estimate the other two parameters with $\mu_\alpha = \mu_\beta = 0.0$. Note that the two described methods to estimate all the parameters start by using good estimates of γ.

5 Numerical Experiments

In all experiments presented in this section, the known part of the PSF was modeled by the Gaussian-shaped PSF defined as

$$\overline{\mathbf{h}}(i,j) \propto \exp\left(-\frac{i^2+j^2}{2\times 3^2}\right) \quad \text{for } i,j=-15,-14,\ldots,-1,0,1,\ldots,14,15, \tag{33}$$

The first of the two algorithms that simultaneously restore the image and estimate the associated hyperparameter uses the fixed-$\mathbf{f}$ covariance model [Eqs. (21)–(23) and (25)] and is named EA1. The second one, using the averaged-$\mathbf{f}$ covariance model [Eqs. (27)–(29) and (31)], is named EA2. In Table 1 some general comments on complexity and speed of the EA1 and EA2 algorithms are shown. Note also that since EA2 approximates $\mathbf{F}\mathbf{F}^t$ by $\alpha\mathbf{Q}^t\mathbf{Q}$, EA1 will, in general, outperform EA2 unless the real underlying image is a "true" realization from the prior model (see Experiment I below).

A number of experiments were performed with the proposed algorithms using a synthetic image obtained from a prior distribution, the "Lena" image, and a real astronomical image to demonstrate their performance. For both algorithms, the criteria $|x^{(n+1)}-x^{(n)}|/x^{(n)} \leqslant 10^{-3}$, $x \in \{\alpha,\beta,\gamma\}$, or $n=250$, whichever was met first, were used to terminate the iterative process. The levels of the PSF and additive noise are measured using the signal-to-noise ratio (SNR), i.e., $\mathrm{SNR}_h = \|\overline{\mathbf{h}}\|^2/N\beta^{-1}$ and $\mathrm{SNR}_g = \|\mathbf{f}\|^2/N\gamma^{-1}$, where $\|\overline{\mathbf{h}}\|^2$ and $\|\mathbf{f}\|^2$ are the energy of the known part of the PSF and of the original image, respectively. The performance of the restoration algorithms was evaluated by measuring the improvement in SNR, denoted by ISNR and given by $\mathrm{ISNR} = \|\mathbf{f}-\mathbf{g}\|^2/\|\mathbf{f}-\hat{\mathbf{f}}\|^2$, where $\mathbf{f}$, $\mathbf{g}$, and $\hat{\mathbf{f}}$ are the original, observed, and estimated images, respectively.

According to Eq. (2), the PSF defined in Eq. (33) was degraded by additive noise in order to obtain a PSF with $\mathrm{SNR}_h = 10$, 20, and 30 dB. Using these degraded PSFs, both synthetic and "Lena" images were blurred, and then Gaussian noise was added to obtain $\mathrm{SNR}_g = 10$, 20, and 30 dB, which produced a set of 18 test images. In all the experiments we assigned to the normalized confidence parameters μ_x, $x \in \{\alpha,\beta,\gamma\}$, the values 0.0,0.1,...,1.0, and we plotted the ISNR as a function of two of the parameters μ_x, $x \in \{\alpha,\beta,\gamma\}$, while keeping the other constant.

Experiment I. To compare the ISNR of the EA1 and EA2 algorithms assuming the means of the hyperpriors are the true values of the noise and image parameters, we used a synthetic image generated from a SAR distribution with $\alpha^{-1} = 2733$.

We have found that the EA1 and EA2 algorithms give very similar results in terms of ISNR on this image, since $\mathbf{f}$ follows a SAR distribution and hence the two models are themselves very similar. Furthermore, we also found that when gamma hyperpriors are used, both methods always gave accurate parameter estimates and the improvement in ISNR is always almost identical. We also noticed that when $\mu_\alpha = \mu_\beta = \mu_\gamma = 0.0$, the estimated value of β was unreliable: it was always very close to zero ($\hat{\beta}$ was between two and three orders of magnitude lower than the true value), and furthermore, for most of the values of SNR_g and SNR_h, both algorithms stopped after reaching the maximum number of iterations (250). The evolution of the ISNR for $\mathrm{SNR}_h = 10$ dB ($\beta^{-1} = 5.6\times 10^{-9}$) and $\mathrm{SNR}_g = 10$ dB ($\gamma^{-1} = 270.85$), for constant $\mu_\beta = 0.0$, is shown in Fig. 1. Similar plots are obtained for experiments with the other SNR_h and SNR_g.

Experiment II. In this experiment both the "Lena" and the SAR image were used. Note that exact knowledge of the parameter α is not possible for the "Lena" image. In order to include some prior knowledge about α and γ, we used the ML estimates[18] for the known PSF problem, assuming that the blurring function was the known part of the PSF. The ML estimates for the "Lena" image with $\mathrm{SNR}_h = 10$ dB and $\mathrm{SNR}_g = 10$ dB were $\alpha^{-1} = 33.1$ and $\gamma^{-1} = 272.77$, and for the SAR image ($\mathrm{SNR}_h = 10$ dB and $\mathrm{SNR}_g = 10$ dB) were $\alpha^{-1} = 165.97$ and $\gamma^{-1} = 270.0$; their corresponding ISNR's were 3.12 and 3.92 dB, respectively.

We observed that including the previously obtained ML estimate as prior knowledge about γ slightly increases the quality of the resulting restoration. However, including the corresponding estimate of the value of α decreases it. This is due to the fact that this ML estimate is accurate for the image noise parameter γ, while it underestimates the variance value α, as shown in Fig. 2.

We also found that, in most cases, EA1 performs a bit better than EA2 algorithm for the "Lena" image. Here

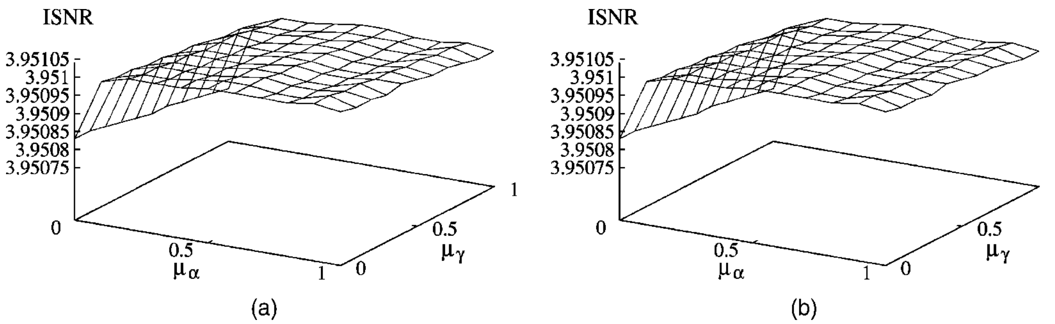


Fig. 1 ISNR evolution with μ_α and μ_γ for $\mu_\beta=0.0$, using the real values of α and γ for the SAR image with $\text{SNR}_h=10$ dB and $\text{SNR}_g=10$ dB, (a) for EA1 algorithm, (b) for EA2 algorithm.

again, when $\mu_\alpha=\mu_\beta=\mu_\gamma=0.0$, both algorithms result in an estimate of β very close to zero ($\hat{\beta}$ was between one and three orders of magnitude less than the real value).

Experiment III. In order to demonstrate that including accurate information about the value of the hyperparameters improves the results, we have tested both algorithms on the "Lena" image, assuming that the means of the hyperpriors for β and γ are the true noise parameters. Since it is not possible to know the real value of α for this image, we used $\mu_\alpha=0.0$, letting the algorithms estimate α without prior information. Figure 3 shows the evolution of the ISNR for $\text{SNR}_h=10$ dB ($\beta^{-1}=5.6\times10^{-9}$) and $\text{SNR}_g=10$ dB ($\gamma^{-1}=273.46$), for constant $\mu_\alpha=0.0$ for the EA1 and EA2 algorithms. Note that incorporating accurate information about the value of β improves more the ISNR than including information about γ, especially for the EA1 algorithm. This is due to the fact that the estimated value of γ is quite close to the real one even if $\mu_\gamma=0.0$, while if no knowledge about the value of β is included, both methods give poor β estimates ($\hat{\beta}$ was between one and three orders of magnitude lower than the real value) when the algorithm stopped, and in most cases the imposed iteration limit was reached. Including knowledge about the real value of β leads to more accurate estimations, since we are forcing

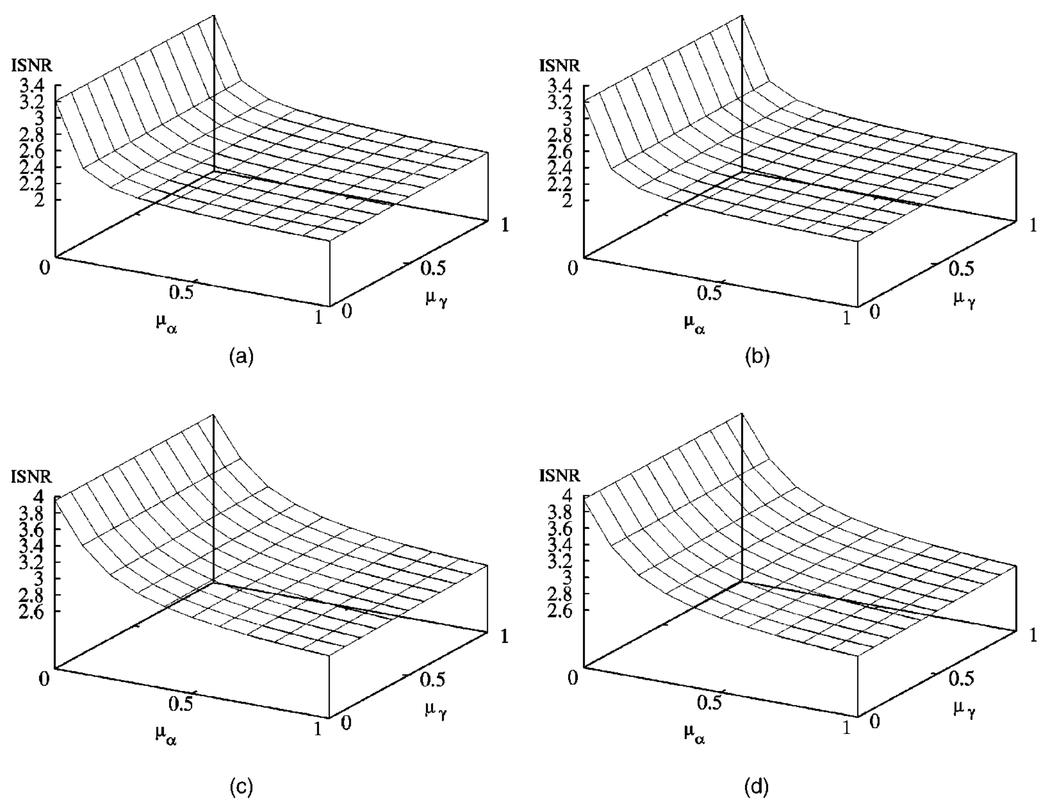


Fig. 2 ISNR evolution with μ_α and μ_γ for $\mu_\beta=0.0$: for (a) EA1 and (b) EA2 algorithms, using the MLE estimated values of α and γ for the "Lena" image with $\text{SNR}_h=10$ dB and $\text{SNR}_g=10$ dB; and for (c) EA1 and (d) EA2 algorithms, using the MLE estimated values of α and γ for the SAR image with $\text{SNR}_h=10$ dB and $\text{SNR}_g=10$ dB.

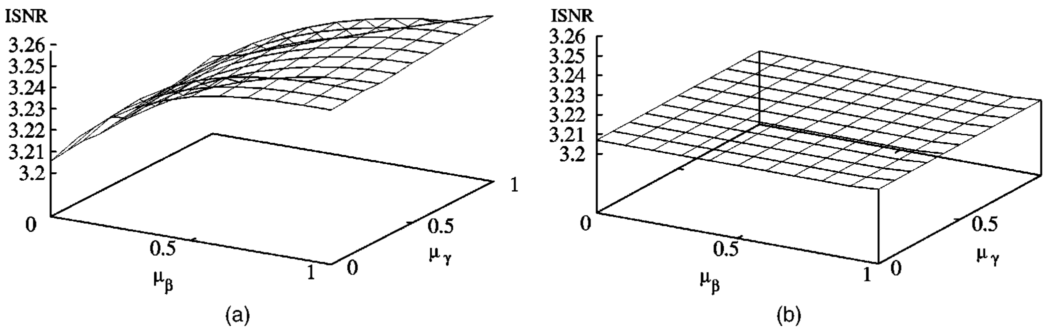


Fig. 3 ISNR evolution with μ_β and μ_γ for $\mu_\alpha=0.0$, using the real values of β and γ for the ''Lena'' image with $\mathrm{SNR}_h=10$ dB and $\mathrm{SNR}_g=10$ dB, (a) for EA1 algorithm, (b) for EA2 algorithm.

Fig. 4 (a) Degraded ''Lena'' image with $\mathrm{SNR}_h=10$ dB and $\mathrm{SNR}_g=10$ dB; (b) restoration with EA1 algorithm, ISNR=3.26 dB; (c) restoration with EA2 algorithm, ISNR=3.21 dB.

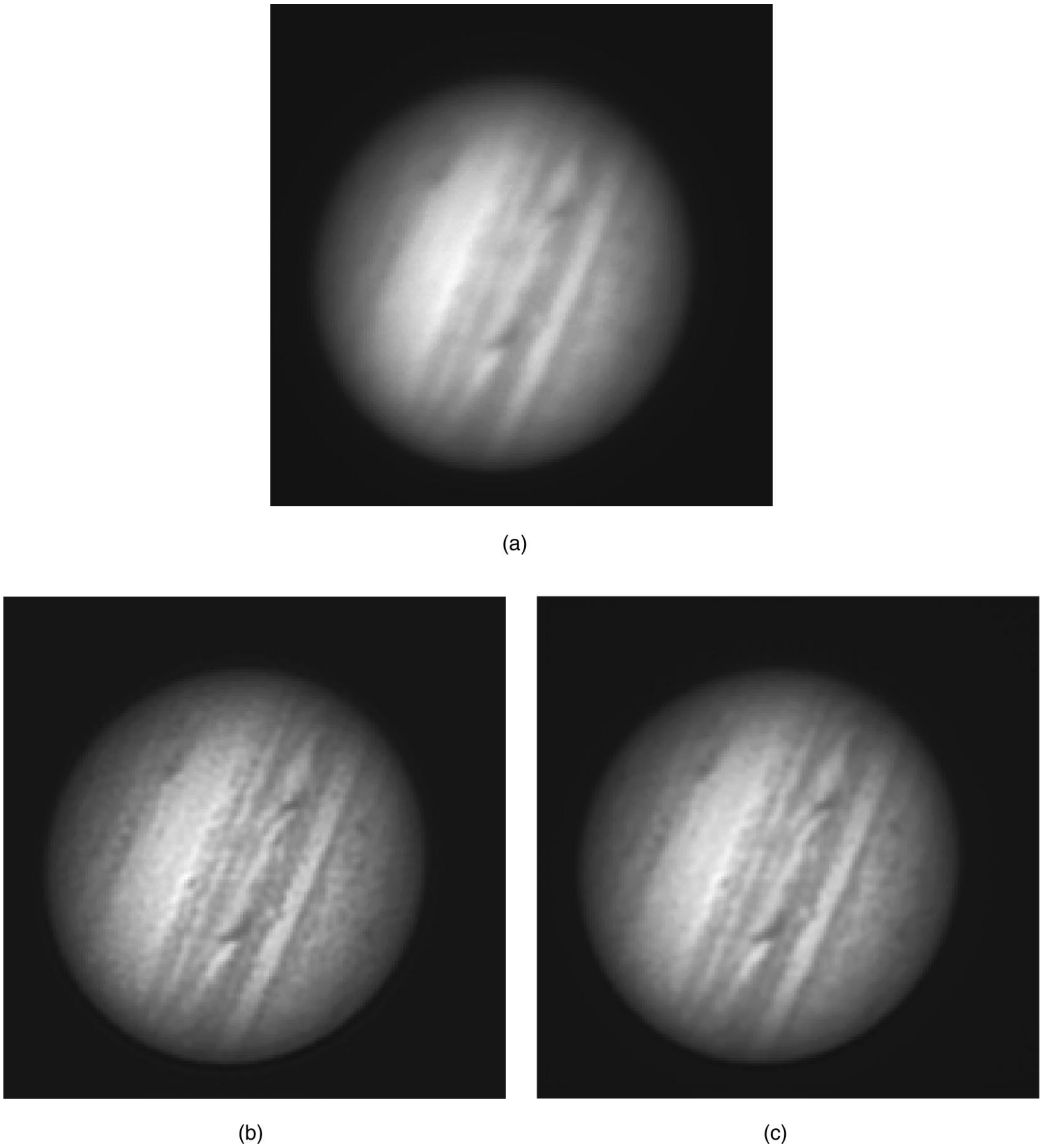


Fig. 5 (a) Observed Jupiter image; (b) restoration with EA1 algorithm; (c) restoration with EA2 algorithm.

both algorithms to provide a β estimate greater than zero. In most cases, also, EA1 performs better than EA2 for this image.

An example of the restoration of the degraded ''Lena'' image [Fig. 4(a), $\mathrm{SNR}_h = 10$ dB, $\mathrm{SNR}_g = 10$ dB] by the EA1 and EA2 algorithms with $\mu_\alpha = 0.0$ and $\mu_\beta = \mu_\gamma = 1.0$ is presented in Figs. 4(b) and 4(c), respectively.

Experiment IV. We also tested the methods on real images. Results are reported on a Jupiter image [depicted in Fig. 5(a)] obtained at Calar Alto Observatory (Spain), using a ground-based telescope, in August 1992. For this kind of images there is no exact expression describing the shape of the PSF, although previous studies[19] have suggested the following radially symmetric approximation for the PSF:

$$h(i,j) \propto \left(1 + \frac{i^2 + j^2}{R^2}\right)^{-B}, \tag{34}$$

where the parameters B and R were estimated from the image to be $B \approx 3$ and $R \approx 5$.[20] However, the estimate of the PSF is not exact, since factors such as atmospheric turbulence introduce noise into it.

Since the proposed methods do not provide reliable estimates simultaneously for both the PSF and additive noise variances, they were estimated in two steps. The algorithms were first run with no prior knowledge about any of the hyperparameters, that is, $\mu_\alpha = \mu_\beta = \mu_\gamma = 0.0$ was used, in order to obtain an estimate of the noise variance. A high confidence was then given to the estimate of γ^{-1}, viz., $\mu_\gamma = 0.8$, and estimates of the other two parameters were obtained. The EA1 and EA2 algorithms were terminated after 46 and 44 iterations, respectively, with the following estimates: $\alpha = 23077.7$, $\beta = 1.27 \times 10^{-10}$, and $\gamma = 47.6$ for the EA1 algorithm, and $\alpha = 22710.1$, $\beta = 1.24 \times 10^{-7}$, and $\gamma = 47.5$ for the EA2 algorithm. The resulting images are shown respectively in Figs. 5(b) and 5(c). It is clear that both algorithms provide good restorations, although the restoration provided by the EA1 algorithm seems to be better resolved.

Alternatively, it is possible to estimate the additive noise variance γ^{-1} using the ML approach as described in Ref. 18, assuming that the PSF is known, as described by Eq. (34). This value is in turn used in the algorithms for the estimation of the remaining parameters. The experimental results provided very similar restorations in the two cases.

6 Conclusions

In this paper we have extended the EA1 and EA2 algorithm and the EM algorithm from our previous work in Refs. 14 and 10, respectively, to include prior knowledge about the unknown parameters. The resulting parameter updates, in both EA1 and EA2 approaches, combine the available prior knowledge with the ML estimates in a simple and intuitive manner. Both algorithms showed the capability to accurately estimate all three parameters simultaneously while restoring the image, even with very low confidence in the prior knowledge. We have also shown that the image noise parameter obtained by the ML estimate for the exactly known PSF problem can be used to guide the estimates of the noise parameter for the partially known PSF problem.

Acknowledgments

This work has been supported by NSF grant MIP-9309910 and by Comisión Nacional de Ciencia y Tecnología under contract TIC2000-1275.

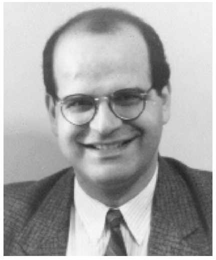
References

1. R. Molina, J. Nuñez, F. Cortijo, and J. Mateos, "Image restoration in astronomy: a Bayesian perspective," *IEEE Signal Process. Mag.* **18**(2), 11–29 (2001).
2. J.-L. Starck, F. Murtagh, and A. Bijaoui, *Image Processing and Data Analysis. The Multiscale Approach*, Cambridge Univ. Press (1998).
3. J. M. Conan, T. Fusco, L. Mugnier, E. Kersal, and V. Michau, "Deconvolution of adaptive optics images with imprecise knowledge of the point spread fonction: results on astonomical objects," in *Astronomy with Adaptive Optics: Present Results and Future Programs*, ESO/ESA (1998).
4. L. Mugnier, J.-M. Conan, T. Fusco, and V. Michau, "Joint maximum a posteriori estimation of object and PSF for turbulence degraded images," in *Bayesian Inference for Inverse Problems*, *Proc. SPIE* **3459**, 50–61 (1998).
5. R. Lagendijk and J. Biemond, *Iterative Identification and Restoration of Images*, Kluwer Academic (1991).
6. L. Guan and R. K. Ward, "Deblurring random time-varying blur," *J. Opt. Soc. Am. A* **6**(11), 1727–1737 (1989).
7. V. Z. Mesarović, N. P. Galatsanos, and A. K. Katsaggelos, "Regularized constrained total least-squares image restoration," *IEEE Trans. Image Process.* **4**(8), 1096–1108 (1995).
8. R. K. Ward and B. E. A. Saleh, "Restoration of images distorted by systems of random impulse response," *J. Opt. Soc. Am. A* **2**(8), 1254–1259 (1995).
9. V. Z. Mesarović, N. P. Galatsanos, and M. N. Wernick, "Restoration from partially-known blur using an expectation-maximization algorithm," in *Proc. Thirtieth ASILOMAR Conf.*, pp. 95–100, Pacific Grove, CA, IEEE (1996).
10. V. N. Mesarović, N. P. Galatsanos, and M. N. Wernick, "Iterative LMMSE restoration of partially-known blurs," *J. Opt. Soc. Am. A* **17**, 711–723 (2000).
11. V. Z. Mesarović, "Image restoration problems under point-spread function uncertainties," PhD Thesis, ECE Dept., Illinois Inst. of Technology, Chicago (1997).
12. V. Z. Mesarović, N. P. Galatsanos, R. Molina, and A. K. Katsaggelos, "Hierarchical Bayesian image restoration from partially-known blurs," in *Proc. Int. Conf. on Acoustics Speech and Signal Processing, ICASSP-98*, Vol. 5, pp. 2905–2908, Seattle, IEEE (1998).
13. N. P. Galatsanos, R. Molina, and V. Z. Mesarović, "Two Bayesian algorithms for image restoration from partially-known blurs," in *Proc. Int. Conf. on Image Processing*, ICIP-98, Vol. 2, pp. 93–97, Chicago, IEEE (1998).
14. N. P. Galatsanos, V. Z. Mesarović, R. Molina, and A. K. Katsaggelos, "Hierarchical Bayesian image restoration from partially-known blurs," *IEEE Trans. Image Process.* **9**(10), 1784–1797 (2000).
15. J. O. Berger, *Statistical Decision Theory and Bayesian Analysis*, Springer-Verlag, New York (1985).
16. B. D. Ripley, *Spatial Statistics*, Wiley (1981).
17. N. P. Galatsanos, V. Z. Mesarović, R. Molina, and A. K. Katsaggelos, "Hyperparameter estimation using hyperpriors for hierarchical Bayesian image restoration from partially-known blurs," in *Bayesian Inference for Inverse Problems*, *Proc. SPIE* **3459**, 337–348 (1998).
18. R. Molina, A. K. Katsaggelos, and J. Mateos, "Bayesian and regularization methods for hyperparameters estimation in image restoration," *IEEE Trans. Image Process.* **8**(2), 231–246 (1999).
19. R. Molina and B. D. Ripley, "Using spatial models as priors in astronomical image analysis," *J. Appl. Statist.* **16**, 193–206 (1989).
20. R. Molina, B. D. Ripley, and F. J. Cortijo, "On the Bayesian deconvolution of planets," in *11th IAPR Int. Conf. on Pattern Recognition*, pp. 147–150 (1992).

Nikolas P. Galatsanos received the Diploma degree in electrical engineering in 1982 from the National Technical University of Athens, Athens, Greece, and the MSEE and PhD degrees in 1984 and 1989, respectively, from the Department of Electrical and Computer Engineering, University of Wisconsin, Madison. He joined the Electrical and Computer Engineering Department, Illinois Institute of Technology, Chicago, in the fall of 1989, where he now holds the rank of associate professor and serves as the director of the graduate program. His current research interests center around image and signal processing and computational vision problems for visual communications and medical imaging applications. He is coeditor of *Recovery Techniques for Image and Video Compression and Transmission* (Kluwer, Norwell, MA, 1998). Dr. Galatsanos served as associate editor of the *IEEE Transactions on Image Processing* from 1992 to 1994. He is an associate editor of the *IEEE Signal Processing Magazine*.

Vladimir Z. Mesarović received the Diploma degree in electrical engineering from the University of Belgrade in 1992, and the MS and PhD degrees in electrical and computer engineering from the Illinois Institute of Technology, Chicago, in 1993 and 1997, respectively. Since January 1997, he has been with the Crystal Audio Division, Cirrus Logic, Austin, TX, where he is currently a project manager in the DSP Systems and Software Group. His research interests are in multimedia signal representations, coding and decoding, transmission, and their efficient DSP implementations.

Rafael Molina received a bachelor's degree in mathematics (statistics) in 1979 and the PhD degree in optimal design in linear models in 1983. He became professor of computer science and artificial intelligence at the University of Granada, Granada, Spain, in 2000, and is currently the dean of the Computer Engineering Faculty. His areas of research interest are image restoration (applications to astronomy and medicine), parameter estimation, image and video compression, and blind deconvolution. Dr. Molina is a member of SPIE, the Royal Statistical Society, and the Asociación Española de Reconocimento de Formas y Análisis de Imágenes (AERFAI).



Aggelos K. Katsaggelos received the Diploma degree in electrical and mechanical engineering from the Aristotelian University of Thessaloniki, Thessaloniki, Greece, in 1979, and the MS and PhD degrees, both in electrical engineering, from the Georgia Institute of Technology, Atlanta, Georgia, in 1981 and 1985, respectively. In 1985 he joined the Department of Electrical Engineering and Computer Science at Northwestern University, Evanston, IL, where he is currently a professor, holding the Ameritech Chair of Information Technology. He is also the director of the Motorola Center for Communications. During the 1986–1987 academic year he was an assistant professor at Polytechnic University, Department of Electrical Engineering and Computer Science, Brooklyn, NY. His current research interests include image and video recovery, video compression, motion estimation, boundary encoding, computational vision, and multimedia signal processing and communications. Dr. Katsaggelos is a Fellow of the IEEE, an Ameritech Fellow, and a member of the Associate Staff, Department of Medicine, at Evanston Hospital. He is an editorial board member of Academic Press, Marcel Dekker (Signal Processing Series), *Applied Signal Processing*, and the *Computer Journal*, and the editor-in-chief of the *IEEE Signal Processing Magazine*. He has served as an associate editor of several journals and is a member of numerous committees and boards. He is the editor of *Digital Image Restoration* (Springer-Verlag, Heidelberg, 1991), coauthor of *Rate-Distortion Based Video Compression* (Kluwer Academic Publishers, 1997), and coeditor of *Recovery Techniques for Image and Video Compression and Transmission*, (Kluwer Academic Publishers, 1998). He has served as the chair and co-chair of several conferences. He is the coholder of eight international patents, and the recipient of the IEEE Third Millennium Medal (2000), the IEEE Signal Processing Society Meritorious Service Award (2001), and an IEEE Signal Processing Society Best Paper Award (2001).

Javier Mateos received the Diploma and MS degrees in computer science from the University of Granada in 1990 and 1991, respectively, and completed his PhD in computer science at the University of Granada in July 1998. In 1992, he became an assistant professor in the Department of Computer Science and Artificial Inteligence of the University of Granada, and he became associate professor in 2001. His research interests include image restoration and image and video recovery and compression. He is a member of the AERFAI (Asociación Española de Reconocimento de Formas y Análisis de Imágenes).