

Hierarchical Bayesian Image Restoration from Partially Known Blurs

Nikolas P. Galatsanos, *Member, IEEE*, Vladimir Z. Mesarović, Rafael Molina, and Aggelos K. Katsaggelos, *Fellow, IEEE*

Abstract—In this paper, we examine the restoration problem when the point-spread function (PSF) of the degradation system is partially known. For this problem, the PSF is assumed to be the sum of a known deterministic and an unknown random component. This problem has been examined before; however, in most previous works the problem of estimating the parameters that define the restoration filters was not addressed. In this paper, two iterative algorithms that simultaneously restore the image and estimate the parameters of the restoration filter are proposed using evidence analysis (EA) within the hierarchical Bayesian framework. We show that the restoration step of the first of these algorithms is in effect almost identical to the regularized constrained total least-squares (RCTLS) filter, while the restoration step of the second is identical to the linear minimum mean square-error (LMMSE) filter for this problem. Therefore, in this paper we provide a solution to the parameter estimation problem of the RCTLS filter. We further provide an alternative approach to the expectation-maximization (EM) framework to derive a parameter estimation algorithm for the LMMSE filter. These iterative algorithms are derived in the discrete Fourier transform (DFT) domain; therefore, they are computationally efficient even for large images. Numerical experiments are presented that test and compare the proposed algorithms.

Index Terms—Blind image restoration, hierarchical Bayesian models, image restoration.

I. INTRODUCTION

TRADITIONALLY, image restoration algorithms have assumed exact knowledge of the blurring operator. In recent years, 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 for example [9]. In most practical applications, the point-spread function (PSF) is neither unknown nor perfectly known. Usually, some information about the PSF is available. However, this information is never exact.

Manuscript received May 26, 1998; revised February 15, 2000. N. P. Galatsanos and V. Z. Mesarović were supported by the National Science Foundation under Grant MIP-9309910 and R. Molina was supported by the Comisión Nacional de Ciencia y Tecnología under Contract TIC97-989. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Patrick L. Combettes.

N. P. Galatsanos is with the Department of Electrical and Computer Engineering, Armour College of Engineering and Science, Illinois Institute of Technology, Chicago, IL 60613 USA (e-mail: npg@ece.iit.edu).

V. Z. Mesarović is with the Crystal Semiconductor Corporation, Austin, TX 78744 USA.

R. Molina is with the Departamento de Ciencias de la Computación, Universidad de Granada, 18071 Granada, Spain.

A. K. Katsaggelos is with the Department of Electrical and Computer Engineering, Northwestern University, Evanston, IL 60208-3118 USA.

Publisher Item Identifier S 1057-7149(00)06131-5.

To our knowledge, restoration of an image when the PSF is modeled by known mean and an additive random component was addressed for the first time in [23]. In this work, the linear minimum mean-square-error (LMMSE) solution was derived for the continuous case only assuming known statistics for the PSF and the additive noises and no experimental results were provided. In [7], [8], [24], and [25], this problem was revisited. In [7], [8], and [24], the LMMSE filter was derived for this problem for the discrete case and numerical experiments were shown. In these works, a white noise model for the PSF and observation errors was used and knowledge of the error statistics was assumed. A difficulty with the LMMSE filter in general is that the signal covariance is not usually known in practice. In [7], the signal covariance is assumed known, but in [8] and [24], an iterative algorithm was proposed in which the current LMMSE signal estimate is used to update the signal covariance estimate. However, the algorithm is ad-hoc and its convergence properties were not analyzed. The expression derived in [24] for the LMMSE filter is not correct. Furthermore, the filter and the estimation algorithm is not derived for the circulant case; thus, it cannot be efficiently calculated for large images. In [25], the Backus-Gilbert method is proposed for this problem. However, the parameter that controls the tradeoff between resolution and noise was not estimated systematically.

In [2] and [22], the problem under consideration was addressed using the theory of projections onto convex sets. However, these convex sets were described by parameters which were assumed known *a priori*. In [15] this problem was addressed using the regularized constrained total least-squares (RCTLS) framework. However, the parameters that define the RCTLS filter were again assumed known *a priori*.

In the classical restoration problem, where the PSF is exactly known, the ratio of the observation noise variance and the smoothness parameter needs to be estimated (see, for example, [5]). This ratio is usually called the regularization parameter and captures the tradeoff between fidelity to the data versus confidence to the prior information used by the restoration filter. A plethora of algorithms has been proposed in the past to estimate this parameter (see, for example, [5]). The estimation of the parameters of the restoration filter in the partially known PSF problem is a much more difficult for the following reasons:

- 1) In the classical restoration problem, there is only one parameter to be estimated and the restoration filter in most cases can be found in closed form; therefore, a subjective trial-and-error search of a one-dimensional (1-D) space can be easily implemented to obtain a good value for the regularization parameter.
- 2) In contrast, for the partially known PSF restoration problem,

there are three parameters to be estimated: the variance of the PSF noise, the variance of the observation noise, and the parameter that captures the image smoothness. Thus, it is impossible to implement a trial-and-error subjective search in a three-dimensional (3-D) space.

In [16] and [17], an iterative algorithm for simultaneous parameter estimation and image restoration based on the EM algorithm [11] was proposed. It was derived under the assumption that the observed data on the image-dependent noise term is Gaussian, and two models for the image prior were used. Due to this assumption, the restoration step of this algorithm is linear and is identical to the LMMSE filter for this problem.

In this paper, we apply evidence analysis (EA) within the hierarchical Bayesian framework to the partially known blur restoration problem [1], [18], [19], [21]. Hierarchical Bayesian analysis is performed in general by combining the prior information and the observed data into what is called the *posterior* distribution, from which all the inferences are made.

The rest of this paper is organized as follows. In Section II, the imaging and image models are discussed. In Section III, the basic philosophy behind EA is briefly presented and its application to the restoration problem from partially known blur is discussed. In Sections IV and V, two EA based algorithm are derived using two models for the conditional distribution of the data. The relationship between the restoration steps of the proposed two EA algorithms is discussed in Section VI. In Section VII, we present numerical experiments which compare the proposed approaches. Finally, in Section VIII we present our conclusions and suggestions for future research.

II. COMPONENTS OF THE HIERARCHICAL MODEL

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

A. Image Priors

A commonly used model for the image prior is based on the stationary Gaussian zero-mean probability density function (PDF), i.e.,

$$P(\mathbf{f}) = [\det(2\pi\mathbf{R}_f)]^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\mathbf{f}^t\mathbf{R}_f^{-1}\mathbf{f}\right\} \quad (1)$$

where \mathbf{R}_f is the $N \times N$ circulant covariance matrix of \mathbf{f} , and $\det(\cdot)$ denotes the determinant of a matrix (see, for example, [10]). In many practical situations, \mathbf{R}_f in (1) may not be available and must be estimated from the blurred and noisy data. However, \mathbf{R}_f may not be identifiable from a single degraded image due to a large number of unknown parameters in \mathbf{R}_f . To avoid this problem, the simultaneously autoregressive (SAR) image models have been proposed, see for example, [18]–[20]. These models can be described by the following conditional PDF:

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

where α is positive unknown parameter that controls the smoothness of the image and $\|\mathbf{Q}\mathbf{f}\|^2$ captures the image

autoregressive model. It is easy to see that the PDFs in (1) and (2) are equivalent when $\mathbf{R}_f^{-1} = \alpha\mathbf{Q}^t\mathbf{Q}$. 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 [12].

B. Observation Model

In [7] and [24], the space-invariant PSF was represented as the sum of a deterministic component and a stochastic component of zero-mean, i.e.,

$$\mathbf{h} = \bar{\mathbf{h}} + \Delta\mathbf{h} \quad (3)$$

where $\bar{\mathbf{h}} \in \mathcal{R}^N$ and $\Delta\mathbf{h} \in \mathcal{R}^N$ are the deterministic (known) and the random (unknown error) components of the PSF, respectively. This is a very general model that attempts to incorporate the random (unknown error) component of the PSF in the restoration algorithm. The unknown component of the PSF is modeled as stationary zero-mean white noise with $N \times N$ covariance matrix $\mathbf{R}_{\Delta h} = (1/\beta)\mathbf{I}$, where $1/\beta$ denotes the variance of the PSF noise and \mathbf{I} is the identity matrix. The observation vector \mathbf{g} is also contaminated by zero-mean additive white noise with $N \times N$ covariance matrix $\mathbf{R}_{\Delta g} = (1/\gamma)\mathbf{I}$, where $1/\gamma$ denotes the variance of the observation noise. Furthermore, the noises in the observed data and the PSF are assumed independent of each other and independent from the source image \mathbf{f} . In this case, the image-degradation can be described by the model [7], [15], [24], [25]

$$\mathbf{g} = \mathbf{H}\mathbf{f} + \Delta\mathbf{g} \quad (4)$$

in which

$$\mathbf{H} = \bar{\mathbf{H}} + \Delta\mathbf{H} \quad (5)$$

and $\mathbf{g}, \mathbf{f}, \Delta\mathbf{g} \in \mathcal{R}^N$ are lexicographically ordered representations of the observed degraded image, the source image, and the additive noise in the observed image, respectively. The matrix $\bar{\mathbf{H}}$ is the known (assumed, estimated or measured) component of the $N \times N$ PSF matrix \mathbf{H} ; $\Delta\mathbf{H}$ is the unknown component of the PSF matrix, generated by $\Delta\mathbf{h}$ defined in (3). Throughout the rest of this paper a circulant approximation of Toeplitz matrices [12] will be used to allow calculations to be performed using the discrete Fourier transform (DFT); thus, $\bar{\mathbf{H}}$ and $\Delta\mathbf{H}$ together with $\mathbf{R}_{\Delta h}, \mathbf{R}_{\Delta g}$ are $N \times N$ circulant matrices [12].

From (3)–(5), it is clear that the form of the conditional distribution of \mathbf{g} is not simple. In what follows we propose two models for $P(\mathbf{g} | \mathbf{f}, \alpha, \beta, \gamma)$.

1) *Fixed- \mathbf{f} Covariance Model*: For this model, we use a Gaussian assumption for both the PSF noise $\Delta\mathbf{h}$ and the additive noise $\Delta\mathbf{g}$ [see (3)–(5)]. Then, to determine $P(\mathbf{g} | \mathbf{f}, \alpha, \beta, \gamma)$ since vector \mathbf{f} is not a random quantity but rather a fixed one, following [14] it is straightforward to see from (4) that $P(\mathbf{g} | \mathbf{f}, \alpha, \beta, \gamma)$ is given by

$$P(\mathbf{g} | \mathbf{f}, \alpha, \beta, \gamma) \propto [\det(\mathbf{R}_{g|\mathbf{f}})]^{-\frac{1}{2}} \times \exp\left\{-\frac{1}{2}(\mathbf{g} - \bar{\mathbf{H}}\mathbf{f})^t\mathbf{R}_{g|\mathbf{f}}^{-1}(\mathbf{g} - \bar{\mathbf{H}}\mathbf{f})\right\}. \quad (6)$$

The conditional covariance $\mathbf{R}_{g|f}$ in (6) is given by

$$\begin{aligned}\mathbf{R}_{g|f} &= E\{(\Delta\mathbf{H}\mathbf{f} + \Delta\mathbf{g})(\Delta\mathbf{H}\mathbf{f} + \Delta\mathbf{g})^t | \mathbf{f}\} \\ &= E\{(\mathbf{F}\Delta\mathbf{h} + \Delta\mathbf{g})(\mathbf{F}\Delta\mathbf{h} + \Delta\mathbf{g})^t | \mathbf{f}\}\end{aligned}\quad (7)$$

where we have used the commutative property of the convolution operation, \mathbf{F} denotes the circulant matrix generated by the image \mathbf{f} , $\Delta\mathbf{h}$ is the unknown PSF noise vector from (3) and $E\{\cdot\}$ denotes the expectation operator. Equation (7) can be further simplified as follows:

$$\begin{aligned}\mathbf{R}_{g|f} &= \mathbf{F}E\{\Delta\mathbf{h}\Delta\mathbf{h}^t\}\mathbf{F}^t + E\{\Delta\mathbf{g}\Delta\mathbf{g}^t\} \\ &= \mathbf{F}\mathbf{R}_{\Delta h}\mathbf{F}^t + \mathbf{R}_{\Delta g} \\ &= \frac{1}{\beta}\mathbf{F}\mathbf{F}^t + \frac{1}{\gamma}\mathbf{I}.\end{aligned}\quad (8)$$

2) *Averaged-f Covariance Model*: For this model, we assume that the *observations* \mathbf{g} are Gaussian and instead of using $\mathbf{F}\mathbf{F}^t$ in the expression for the covariance we use its mean value from the prior. In other words, the expectation in (7) is taken over \mathbf{f} also. Thus, for this model we assume a Gaussian $P(\mathbf{g} | \mathbf{f}, \alpha, \beta, \gamma)$ with mean $\bar{\mathbf{H}}\mathbf{f}$ and covariance $\underline{\mathbf{R}}_{g|f}$ given by

$$\underline{\mathbf{R}}_{g|f} = \left[\frac{N}{\beta}(\alpha\mathbf{Q}^t\mathbf{Q})^{-1} + \frac{\mathbf{I}}{\gamma} \right]. \quad (9)$$

Note that by using this approximation we have incorporated the uncertainty of the image prior model, α , in the conditional distribution. Thus, we 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 what follows.

Because of the attractiveness of the Bayesian formulation, it is convenient to use priors on the hyperparameters even when very little prior information is available about certain parameters [1]. According to [1], in situations where no prior information is available a noninformative prior is used. This applies to the hyperparameters $\omega \in \{\alpha, \beta, \gamma\}$ that were previously introduced. In this paper we use improper, non informative priors $p(\omega) \propto \text{const}$ over $[0, \infty)$.

III. HIERARCHICAL BAYESIAN ANALYSIS

Let us examine how the hierarchical Bayesian analysis is performed on our partially known blur restoration problem. After defining $P(\alpha, \beta, \gamma, \mathbf{f}, \mathbf{g})$, the Bayesian analysis can be carried out in two different ways. In the evidence analysis (EA) framework, $P(\alpha, \beta, \gamma, \mathbf{f}, \mathbf{g})$ is integrated over \mathbf{f} to give the evidence $P(\alpha, \beta, \gamma | \mathbf{g})$ which is then maximized over the hyper-parameters; the restoration is then performed using the estimated hyper-parameters. In the MAP framework $P(\alpha, \beta, \gamma, \mathbf{f}, \mathbf{g})$ is integrated over α, β and γ to obtain the true likelihood which is then maximized with respect to \mathbf{f} . In this work we shall use EA instead of MAP analysis. We have found that the EA formulation provides better results since it allows to estimate the hyper-parameters for restoration–reconstruction problems. In [13] and [19], a detailed discussion is provided of the merits of EA over MAP for these problems.

According to the EA approach, the simultaneous estimation of $\mathbf{f}, \alpha, \beta$, and γ is done as follows.

Parameter estimation step:

$$\hat{\alpha}, \hat{\beta}, \hat{\gamma} = \arg \max_{\alpha, \beta, \gamma} \{P(\alpha, \beta, \gamma | \mathbf{g})\}. \quad (10)$$

Restoration step:

$$\hat{\mathbf{f}}(\hat{\alpha}, \hat{\beta}, \hat{\gamma}) = \arg \max_{\mathbf{f}} \{P(\mathbf{f} | \mathbf{g}, \hat{\alpha}, \hat{\beta}, \hat{\gamma})\}. \quad (11)$$

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 current estimates of the parameters. Therefore, the above two-step procedure is repeated until convergence occurs.

In order to find $P(\alpha, \beta, \gamma | \mathbf{g})$, as required by the parameter estimation step, we take into account that from the distributions defined in Section II we have

$$\begin{aligned}P(\mathbf{g}, \mathbf{f}, \alpha, \beta, \gamma) \\ = P(\mathbf{g} | \mathbf{f}, \alpha, \beta, \gamma)P(\mathbf{f} | \alpha, \beta, \gamma)P(\alpha)P(\beta)P(\gamma).\end{aligned}\quad (12)$$

Then, to obtain $P(\alpha, \beta, \gamma | \mathbf{g})$ as required in (10) we marginalize the PDF in (12) with respect to \mathbf{f} [1], [18], and [19], i.e.,

$$P(\alpha, \beta, \gamma | \mathbf{g}) \propto \int P(\mathbf{g}, \mathbf{f}, \alpha, \beta, \gamma) d\mathbf{f}. \quad (13)$$

Since we assumed “flat” noninformative hyper-priors, $P(\alpha)P(\beta)P(\gamma)$ can be discarded in (12) and so we have

$$P(\mathbf{g}, \mathbf{f}, \alpha, \beta, \gamma) \propto P(\mathbf{f} | \alpha, \beta, \gamma)P(\mathbf{g} | \mathbf{f}, \alpha, \beta, \gamma). \quad (14)$$

The use of Gamma hyper-priors for this problem is described in [6] and [17].

Now, as required in (11) for the restoration step, the image *posterior* PDF $P(\mathbf{f} | \mathbf{g}, \hat{\alpha}, \hat{\beta}, \hat{\gamma})$ can be obtained applying Bayes rule to the joint PDF, i.e.,

$$P(\mathbf{f} | \mathbf{g}, \hat{\alpha}, \hat{\beta}, \hat{\gamma}) \propto P(\mathbf{f} | \hat{\alpha}, \hat{\beta}, \hat{\gamma})P(\mathbf{g} | \mathbf{f}, \hat{\alpha}, \hat{\beta}, \hat{\gamma}) \quad (15)$$

where $P(\mathbf{f} | \hat{\alpha}, \hat{\beta}, \hat{\gamma})$ and $P(\mathbf{g} | \mathbf{f}, \hat{\alpha}, \hat{\beta}, \hat{\gamma})$ are given in (2) and (6), respectively, evaluated at $\hat{\alpha}, \hat{\beta}$, and $\hat{\gamma}$.

Using the two different choices for the conditional covariance given in given in (8) and (9), we will now proceed with the evidence analysis.

IV. EVIDENCE ANALYSIS BASED ON THE FIXED- \mathbf{f} COVARIANCE MODEL

Substituting (2) and (6) into (12) we obtain

$$\begin{aligned}P(\mathbf{g}, \mathbf{f}, \alpha, \beta, \gamma) &\propto \alpha^{\left(\frac{N-1}{2}\right)} [\det(\mathbf{R}_{g|f})]^{-\frac{1}{2}} \\ &\times \exp\left\{-\frac{1}{2}J(\mathbf{f}, \alpha, \beta, \gamma)\right\}\end{aligned}\quad (16)$$

where

$$J(\mathbf{f}, \alpha, \beta, \gamma) = \alpha \|\mathbf{Q}\mathbf{f}\|^2 + (\mathbf{g} - \bar{\mathbf{H}}\mathbf{f})^t \mathbf{R}_{g|f}^{-1} (\mathbf{g} - \bar{\mathbf{H}}\mathbf{f}). \quad (17)$$

A. Parameter Estimation Step

To compute $P(\alpha, \beta, \gamma | \mathbf{g})$ as required by the parameter estimation step, we substitute (16) into (13). This gives

$$P(\alpha, \beta, \gamma | \mathbf{g}) \propto \alpha^{\frac{N-1}{2}} \int [\det(\mathbf{R}_{g|f})]^{-\frac{1}{2}} \times \exp\left\{-\frac{1}{2}J(\mathbf{f}, \alpha, \beta, \gamma)\right\} d\mathbf{f}. \quad (18)$$

Now, we are ready to perform the integration in (18). First, 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.,

$$J(\mathbf{f}, \alpha, \beta, \gamma) \approx J(\mathbf{f}^{(n)}, \alpha, \beta, \gamma) + (\mathbf{f} - \mathbf{f}^{(n)})^t \nabla J(\mathbf{f}, \alpha, \beta, \gamma)|_{\mathbf{f}^{(n)}} + \frac{1}{2} (\mathbf{f} - \mathbf{f}^{(n)})^t \nabla^2 J(\mathbf{f}, \alpha, \beta, \gamma)|_{\mathbf{f}^{(n)}} (\mathbf{f} - \mathbf{f}^{(n)}). \quad (19)$$

Next, we observe that in (19)

$$\nabla J(\mathbf{f}, \alpha, \beta, \gamma)|_{\mathbf{f}^{(n)}} = 0 \quad (20)$$

if $\mathbf{f}^{(n)}$ is chosen to be the minimizer of $J(\mathbf{f}, \alpha, \beta, \gamma)$ in (17), and that the Hessian matrix can be approximated by

$$\nabla^2 J(\mathbf{f}, \alpha, \beta, \gamma)|_{\mathbf{f}^{(n)}} = \mathbf{G}^{(n)} = \alpha \mathbf{Q}^t \mathbf{Q} + \bar{\mathbf{H}}^t \mathbf{R}_{g|f^{(n)}}^{-1} \bar{\mathbf{H}} \quad (21)$$

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

Finally, substituting (19) into (18), and by using the fact that $[\det(\mathbf{R}_{g|f})]^{-(1/2)}$ depends on \mathbf{f} weakly, compared to the exponential term under the integral and so it can be substituted by $[\det(\mathbf{R}_{g|f^{(n)}})]^{-\frac{1}{2}}$ and that integral of a PDF is equal to 1, (18) becomes

$$P(\alpha, \beta, \gamma | \mathbf{g}) \propto \alpha^{\frac{N-1}{2}} \det[\mathbf{R}_{g|f^{(n)}}]^{-\frac{1}{2}} \det[\mathbf{G}^{(n)}]^{-\frac{1}{2}} \times \exp\left\{-\frac{1}{2}J(\mathbf{f}^{(n)}, \alpha, \beta, \gamma)\right\}. \quad (22)$$

It is interesting to note that the *posterior* functional in (22) is equivalent to the likelihood functional $P(\mathbf{g} | \mathbf{f}, \alpha, \beta, \gamma)$ for this problem since we assumed uniform priors on α, β , and γ .

In the above equations we have used the fact that $[\det(\mathbf{R}_{g|f})]^{-\frac{1}{2}}$ depends on \mathbf{f} weakly, compared to the exponential term under the integral. To justify this we observe that the eigenvalues of $\mathbf{R}_{g|f}$ in (8) are given by

$$\frac{1}{\beta} |F(i)|^2 + \frac{1}{\gamma} = \frac{1}{\beta} N S_f(i) + \frac{1}{\gamma} \quad (23)$$

where $F(i)$ is the i th DFT coefficient of the image \mathbf{f} and $S_f(i)$ is the periodogram estimate of the i th power spectrum coefficient of the image \mathbf{f} (the i th eigenvalue of the image covariance matrix). The power spectrum (or the covariance matrix), however, is a statistic of \mathbf{f} ; it depends on the class of images \mathbf{f} belongs to, not on \mathbf{f} . The same reasons justify the approximation used for the Hessian.

Taking “ $-2 \log$ ” of both sides of (22) we obtain the following functional:

$$L(\alpha, \beta, \gamma) = -(N-1) \log \alpha + \log \det[\mathbf{R}_{g|f^{(n)}}] + \log \det[\mathbf{G}^{(n)}] + J(\mathbf{f}^{(n)}, \alpha, \beta, \gamma) \quad (24)$$

which has to be minimized.

To minimize this functional we can use the following iterative scheme whose complete derivation can be found in Appendix A

$$\frac{1}{\alpha^{(n+1)}} = \left[\|\mathbf{Q}\mathbf{f}^{(n)}\|^2 + \text{tr}[\mathbf{G}^{(n)-1} \mathbf{Q}^t \mathbf{Q}] \right] / (N-1) \quad (25)$$

$$\begin{aligned} \frac{1}{\beta^{(n+1)}} &= \left[\text{tr}[\mathbf{R}_{g|f^{(n)}}^{-1}] \right] / [N\beta^{(n)}\gamma^{(n)}] \\ &+ \left[\text{tr}[\mathbf{G}^{(n)-1} \bar{\mathbf{H}}^t \mathbf{R}_{g|f^{(n)}}^{-2} \mathbf{F}^{(n)} \mathbf{F}^{(n)t} \bar{\mathbf{H}}] \right. \\ &+ \left. (\mathbf{g} - \bar{\mathbf{H}}\mathbf{f}^{(n)})^t \mathbf{R}_{g|f^{(n)}}^{-2} \mathbf{F}^{(n)} \mathbf{F}^{(n)t} \right. \\ &\times \left. (\mathbf{g} - \bar{\mathbf{H}}\mathbf{f}^{(n)}) \right] / [N\beta^{(n)2}] \end{aligned} \quad (26)$$

$$\begin{aligned} \frac{1}{\gamma^{(n+1)}} &= \left[\text{tr}[\mathbf{F}^{(n)} \mathbf{F}^{(n)t} \mathbf{R}_{g|f^{(n)}}^{-1}] \right] / [N\beta^{(n)}\gamma^{(n)}] \\ &+ \left[\text{tr}[\mathbf{G}^{(n)-1} \bar{\mathbf{H}}^t \mathbf{R}_{g|f^{(n)}}^{-2} \bar{\mathbf{H}}] \right. \\ &+ \left. (\mathbf{g} - \bar{\mathbf{H}}\mathbf{f}^{(n)})^t \mathbf{R}_{g|f^{(n)}}^{-2} (\mathbf{g} - \bar{\mathbf{H}}\mathbf{f}^{(n)}) \right] / [N\gamma^{(n)2}] \end{aligned} \quad (27)$$

where $\mathbf{f}^{(n)}, \mathbf{G}^{(n)}, \mathbf{F}^{(n)}, \alpha^{(n)}, \beta^{(n)}, \gamma^{(n)}$ are calculated at iteration (n) . It is important to note that this iterative scheme can also be carried out in the Fourier domain (see Appendix B).

The parameter estimation cycle in (25)–(27) is repeated until convergence in (24) occurs. Although the proof of convergence of the resulting parameter estimators seems to be analytically intractable, in all our experiments with this EA algorithm we observed not only convergence in the *posterior* functional, but also in terms of the parameter values.

B. Restoration Step

To perform the restoration step, we take into account that

$$\arg \max_{\mathbf{f}} \{P(\mathbf{f} | \mathbf{g}, \alpha, \beta, \gamma)\} = \arg \max_{\mathbf{f}} \{P(\mathbf{f}, \mathbf{g}, \alpha, \beta, \gamma)\}. \quad (28)$$

As a result substituting (16) and (17) into (28) and taking the negative log

$$\begin{aligned} \hat{\mathbf{f}}(\hat{\alpha}, \hat{\beta}, \hat{\gamma}) &= \arg \min_{\mathbf{f}} \left\{ (\bar{\mathbf{H}}\mathbf{f} - \mathbf{g})^t \hat{\mathbf{R}}_{g|f}^{-1} (\bar{\mathbf{H}}\mathbf{f} - \mathbf{g}) \right. \\ &\left. + \hat{\alpha} \|\mathbf{Q}\mathbf{f}\|^2 + \log \left[\det(\hat{\mathbf{R}}_{g|f}) \right] \right\} \end{aligned} \quad (29)$$

where $\hat{\mathbf{R}}_{g|f} = (1/\hat{\beta})\mathbf{F}\mathbf{F}^t + (1/\hat{\gamma})\mathbf{I}$. The functional in (29) is nonconvex and may have several local minima. In general, a closed form solution to (29) does not exist and numerical optimization algorithms must be used. A practical computation of (29) can be obtained by transforming it to the DFT domain. In Appendix B we show using the diagonalization properties of the

DFT that minimization of (29) can be performed in the DFT domain as follows:

$$\hat{F}(i) = \arg \min_{F(i)} \left\{ \frac{1}{N} \left[\frac{|\bar{H}(i)F(i) - G(i)|^2}{\frac{1}{\beta}|F(i)|^2 + \frac{1}{\gamma}} + \hat{\alpha}|Q(i)|^2|F(i)|^2 \right] + \log \left[\frac{1}{\beta}|F(i)|^2 + \frac{1}{\gamma} \right] \right\} \quad (30)$$

for each frequency $i = 0, 1, \dots, N-1$. In (30), $G(i)$ and $F(i)$ are the DFT coefficients of the observed and restored images, $\bar{H}(i)$ and $Q(i)$ are the eigenvalues of $\bar{\mathbf{H}}$ and \mathbf{Q} , $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\gamma}$ are the estimates of the hyper-parameters obtained in the parameter estimation step and $|\cdot|$ denotes the modulus of a complex quantity.

The “log” term in (30) is weakly dependent on \mathbf{f} and can be discarded when optimizing with respect to \mathbf{f} . We verified this experimentally by comparing the solutions of (30) with and without the “log” term. We used a number of choices for the initial points for our optimization algorithm (the source, the degraded, and the EM-restored images) to test the point of convergence. In all cases we found that the selection of the initial points did not alter the solution of (30), regardless of whether the “log” term was present. Thus, for all practical purposes in the restoration step we minimize $J(\mathbf{f}, \alpha, \beta, \gamma)$ with respect to \mathbf{f} .

C. Comparison with RCTLS Image Restoration

It was shown in [15] that the RCTLS estimate of the DFT of \mathbf{f} , $F(i)$, is found by minimizing the following function: $J(F(i))$ for $i = 0, 1, \dots, N-1$

$$J(F(i)) = \frac{|\bar{H}(i)F(i) - G(i)|^2}{\frac{1}{\beta}|F(i)|^2 + \frac{1}{\gamma}} + \alpha|Q(i)|^2|F(i)|^2. \quad (31)$$

From (31), it is clear that when the variance of the noise in the PSF $1/\beta$ becomes zero, the RCTLS estimate degenerates, as expected, to the RLS estimate [5]. In that case, the PSF matrix $\bar{\mathbf{H}}$ coincides with \mathbf{H} and the regularization parameter α_{RLS} becomes equal to $(1/\gamma)\alpha_{\text{RCTLS}}$. An open problem with the RCTLS estimator is how to estimate the parameters α, β , and γ . However, (31) and (30) are very similar and in fact the “log” term in (31) is, as already mentioned, discarded when optimizing. So, the proposed EA provides, together with an alternative interpretation of the RCTLS method, an algorithm to estimate the unknown hyper parameters.

V. EVIDENCE ANALYSIS BASED ON THE AVERAGED-F COVARIANCE MODEL

We follow identical steps as in the previous section with

$$\mathbf{R}_{g|f} = \frac{N}{\beta}(\alpha\mathbf{Q}^t\mathbf{Q})^{-1} + \frac{\mathbf{I}}{\gamma}. \quad (32)$$

Taking “ $-2\log$ ” of the corresponding $P(\alpha, \beta, \gamma | \mathbf{g})$ yields a similar form to (24) log-likelihood function given by

$$\begin{aligned} \underline{L}(\alpha, \beta, \gamma) &= -(N-1)\log\alpha + \log\det[\mathbf{R}_{g|f}] \\ &+ \log\det[\mathbf{G}] + \underline{J}(\mathbf{f}^{(n)}, \alpha, \beta, \gamma) \end{aligned} \quad (33)$$

where

$$\underline{J}(\mathbf{f}^{(n)}, \alpha, \beta, \gamma) = \alpha \left\| \mathbf{Q}\mathbf{f}^{(n)} \right\|^2 + \left(\mathbf{g} - \bar{\mathbf{H}}\mathbf{f}^{(n)} \right)^t \times \mathbf{R}_{g|f}^{-1} \left(\mathbf{g} - \bar{\mathbf{H}}\mathbf{f}^{(n)} \right) \quad (34)$$

and

$$\nabla^2 \tilde{J}(\mathbf{f}, \alpha, \beta, \gamma)|_{\mathbf{f}^{(n)}} = \mathbf{G} = \alpha\mathbf{Q}^t\mathbf{Q} + \bar{\mathbf{H}}^t\mathbf{R}_{g|f}^{-1}\bar{\mathbf{H}}. \quad (35)$$

Clearly, the functional relationship of β and γ remains the same in the two likelihood functions $L(\alpha, \beta, \gamma)$ in (24) and $\underline{L}(\alpha, \beta, \gamma)$ in (33). However, the functional relationship of α changes in (24) and (33). Therefore, we expect the update equations for β and γ to be very similar to the ones of the previous algorithm but not the update for α . Furthermore, since $\mathbf{R}_{g|f}$ does not depend on \mathbf{f} , $\underline{J}(\mathbf{f}^{(n)}, \alpha, \beta, \gamma)$ is quadratic with respect to \mathbf{f} . As a result, the image restoration step gives a linear estimate for \mathbf{f} for this algorithm.

A. Parameter Estimation Step

To find the estimates of the parameters $\underline{L}(\alpha, \beta, \gamma)$ must be minimized. Taking the derivatives of $\underline{L}(\alpha, \beta, \gamma)$ with respect to α, β, γ , setting them equal to zero and taking these equations to the DFT domain, yields the following iterations for α, β , and γ :

$$\begin{aligned} \frac{N-1}{\alpha^{(n+1)}} &= \sum_{i=0}^{N-1} |Q(i)|^2 \left(S^{(n)}(i) + \frac{1}{N} |F^{(n)}(i)|^2 \right) \\ &- \sum_{i=0}^{N-1} \frac{N \frac{1}{\beta^{(n)}}}{N\alpha^{(n)} \frac{1}{\beta^{(n)}} + \alpha^{(n)2} |Q(i)|^2 \frac{1}{\gamma^{(n)}}} \\ &+ \sum_{i=0}^{N-1} \frac{N \frac{1}{\beta^{(n)}}}{\alpha^{(n)2} |Q(i)|^2 \left[\frac{N}{\alpha^{(n)}\beta^{(n)}|Q(i)|^2} + \frac{1}{\gamma^{(n)}} \right]^2} \\ &\times \left\{ |\bar{H}(i)|^2 S^{(n)}(i) + \frac{1}{N} |\bar{H}(i)F^{(n)}(i) - G(i)|^2 \right\} \end{aligned} \quad (36)$$

$$\begin{aligned} \frac{1}{\beta^{(n+1)}} &= \frac{1}{N} \sum_{i=0}^{N-1} \frac{1}{\beta^{(n)}\gamma^{(n)}} \frac{1}{N \frac{1}{\beta^{(n)}} \frac{1}{\alpha^{(n)}|Q(i)|^2} + \frac{1}{\gamma^{(n)}}} \\ &+ \frac{1}{N} \sum_{i=0}^{N-1} \frac{1}{\beta^{(n)2}} \frac{N \frac{1}{\alpha^{(n)}|Q(i)|^2}}{\left[N \frac{1}{\beta^{(n)}} \frac{1}{\alpha^{(n)}|Q(i)|^2} + \frac{1}{\gamma^{(n)}} \right]^2} \\ &\times \left\{ |\bar{H}(i)|^2 S^{(n)}(i) + \frac{1}{N} |\bar{H}(i)F^{(n)}(i) - G(i)|^2 \right\} \end{aligned} \quad (37)$$

$$\begin{aligned} \frac{1}{\gamma^{(n+1)}} &= \frac{1}{N} \sum_{i=0}^{N-1} \frac{1}{\beta^{(n)}\gamma^{(n)}} \frac{N \frac{1}{\alpha^{(n)}|Q(i)|^2}}{N \frac{1}{\beta^{(n)}} \frac{1}{\alpha^{(n)}|Q(i)|^2} + \frac{1}{\gamma^{(n)}}} \\ &+ \frac{1}{N} \sum_{i=0}^{N-1} \frac{1}{\gamma^{(n)2}} \frac{1}{\left[N \frac{1}{\beta^{(n)}} \frac{1}{\alpha^{(n)}|Q(i)|^2} + \frac{1}{\gamma^{(n)}} \right]^2} \\ &\times \left\{ |\bar{H}(i)|^2 S^{(n)}(i) + \frac{1}{N} |\bar{H}(i)F^{(n)}(i) - G(i)|^2 \right\} \end{aligned} \quad (38)$$

with

$$S^{(n)}(i) = \frac{S_f^{(n)}(i) \left[N \frac{1}{\beta^{(n)}} \frac{1}{\alpha^{(n)}|Q(i)|^2} + \frac{1}{\gamma^{(n)}} \right]}{\left(|\bar{H}(i)|^2 + N \frac{1}{\beta^{(n)}} \right) \frac{1}{\alpha^{(n)}|Q(i)|^2} + \frac{1}{\gamma^{(n)}}}. \quad (39)$$

The above algorithm is identical to the EM-based algorithm derived for this problem in [16] and [17]. To derive the EM algorithm a SAR prior image model was used for the image prior and $\mathbf{z} = [\mathbf{f}^T, \mathbf{g}^T]^T$ and \mathbf{g} were the complete and incomplete data, respectively. This is no surprise since both the EM and the proposed algorithm find ML estimates for the parameters.

B. Image Restoration Step

For the image estimation step we can write similarly to (29)

$$\begin{aligned} \hat{\mathbf{f}}(\hat{\alpha}, \hat{\beta}, \hat{\gamma}) &= \arg \min_{\mathbf{f}} \{ \mathcal{J}(\mathbf{f}, \hat{\alpha}, \hat{\beta}, \hat{\gamma}) \} \\ &= \arg \min_{\mathbf{f}} \left\{ (\bar{\mathbf{H}}\mathbf{f} - \mathbf{g})^t \underline{\mathbf{R}}_{g|f}^{-1} (\bar{\mathbf{H}}\mathbf{f} - \mathbf{g}) \right. \\ &\quad \left. + \hat{\alpha} \|\mathbf{Q}\mathbf{f}\|^2 + \log[\det(\underline{\mathbf{R}}_{g|f})] \right\} \end{aligned} \quad (40)$$

where $\underline{\mathbf{R}}_{g|f} = (N/\hat{\beta})[\hat{\alpha}\mathbf{Q}^t\mathbf{Q}]^{-1} + (1/\hat{\gamma})\mathbf{I}$. This minimization yields

$$\begin{aligned} \hat{\mathbf{f}}(\hat{\alpha}, \hat{\beta}, \hat{\gamma}) &= \left[\bar{\mathbf{H}}^t(\underline{\mathbf{R}}_{g|f})^{-1}\bar{\mathbf{H}} + \hat{\alpha}\mathbf{Q}^t\mathbf{Q} \right]^{-1} \\ &\quad \times \bar{\mathbf{H}}^t(\underline{\mathbf{R}}_{g|f})^{-1}\mathbf{g}. \end{aligned} \quad (41)$$

Taking (41) in the DFT domain we get

$$\begin{aligned} \hat{F}(i) &= \frac{\bar{H}^*(i) \frac{1}{\hat{\alpha}|\mathbf{Q}(i)|^2}}{\left(|\bar{H}(i)|^2 + \frac{N}{\hat{\beta}} \right) \frac{1}{\hat{\alpha}|\mathbf{Q}(i)|^2} + \frac{1}{\hat{\gamma}}} G(i), \\ &\quad \text{for } i = 0, 1, \dots, N-1 \end{aligned} \quad (42)$$

where $Q(i)$ and $\bar{H}(i)$ are the eigenvalues of \mathbf{Q} and $\bar{\mathbf{H}}$, respectively, and $G(i)$ is the DFT of \mathbf{g} .

C. Relation to LMMSE Image Restoration

The LMMSE filter for this problem [7], [24] is given by

$$\hat{\mathbf{f}} = \mathbf{L}\mathbf{g} = \mathbf{R}_f \bar{\mathbf{H}}^t [\bar{\mathbf{H}}\mathbf{R}_f \bar{\mathbf{H}}^t + \mathbf{R}_T(\mathbf{f})]^{-1} \mathbf{g} \quad (43)$$

where \mathbf{R}_f and $\mathbf{R}_T(\mathbf{f})$ are the covariances of the signal \mathbf{f} and the combined PSF and additive noise $\Delta\mathbf{H}\mathbf{f} + \Delta\mathbf{g}$, respectively.

A few comments are in order regarding the LMMSE estimator in (43). When the PSF and the observation noise are white with variances $1/\beta$ and $1/\gamma$, as assumed in Section II, it was shown in [16], [17] that the LMMSE filter in the DFT domain assumes the form

$$\begin{aligned} \hat{F}(i) &= \frac{S_f(i)H^*(i)}{\left[|H(i)|^2 + \frac{N}{\beta} \right] S_f(i) + \frac{1}{\gamma}} \cdot G(i), \\ &\quad \text{for } i = 0, 1, \dots, N-1. \end{aligned} \quad (44)$$

Thus, (42) is identical to the DFT expression of the LMMSE estimate for this problem in (44) with a SAR image prior ($S_f(i) = (1/\alpha|Q(i)|^2)$).

VI. COMPARISON OF THE RESTORATION STEPS OF THE TWO PROPOSED ALGORITHMS

It is a well-known fact that the linear minimum mean square-error (LMMSE) estimator is identical to the MAP solution under a linear observation model and a Gaussian assumption for the

source signal and the noise [11]. However, in this problem, the observation model in (4) is not linear, owing to the signal-dependent noise term $\Delta\mathbf{H}\mathbf{f} + \Delta\mathbf{g}$. Furthermore, the product $\Delta\mathbf{H}\mathbf{f}$ is not Gaussian if both \mathbf{f} and $\Delta\mathbf{H}$ are Gaussian. Therefore, the LMMSE and the MAP approaches in our problem do not necessarily yield the same solution. The restoration step of the first evidence analysis based algorithm is the MAP estimate for this problem in the ‘‘classical’’ sense since \mathbf{f} is assumed fixed in the conditional of (6). However, the restoration step of the EA2 algorithm, where the conditional covariance $\underline{\mathbf{R}}_{g|f}$ is assumed to be given by (9), is also a MAP estimate. This conditional in effect *linearizes* the observation model and then the MAP and the LMMSE estimates become identical.

In an attempt to derive rigorously the mathematical relationship between the restoration steps of the two evidence analysis based algorithms we perform a perturbation analysis in Appendix C, under the small noise assumption. There we show that the image estimate obtained by the restoration step of the second algorithm is identical to the ‘‘linear’’ approximation of the image estimate obtained in the restoration step of the first algorithm. This result is valid for small noise perturbations only. For comparisons at moderate and high noise levels we resort to computer simulations in the experimental section.

VII. NUMERICAL EXPERIMENTS

In this section, we test and compare the two previously derived EA algorithms. For brevity we shall refer to the first one which is based on the fixed- \mathbf{f} covariance prior model as EA1 and to the second one which is based on the averaged- \mathbf{f} covariance model as EA2. These algorithms are tested under white and correlated PSF-noise perturbations. First, the restoration capabilities of the proposed algorithms are compared using synthetic images that satisfy exactly the SAR image model. Second, these algorithms are compared with natural images for simultaneous parameter estimation and restoration.

The (per pixel) MSE is defined as

$$\text{MSE} = \frac{1}{N} \|\mathbf{f} - \hat{\mathbf{f}}\|_2^2 \quad (45)$$

where \mathbf{f} and $\hat{\mathbf{f}}$ are the original and the restored (upon convergence) images, respectively.

To obtain statistically meaningful results we performed Monte-Carlo simulations in which the MSE was averaged over five different noise realizations. We experimentally observed that more than five noise realizations does not change the nature of the MSE curves.

The MSE is a function of two noise parameters: $1/\beta$ and $1/\gamma$. In order to enhance the clarity and the visibility of the results we plot instead of 3-D plot of the MSE versus both noise parameters two representative 2-D MSE plots: (a) For a fixed $\text{SNR}_g = 30$ dB we plot MSE versus SNR_h by varying $1/\beta$, and label this plot as Plot-H. (b) For a fixed $\text{SNR}_h = 20$ dB we plot MSE versus SNR_g by varying $1/\gamma$, and label this plot as Plot-G. In these plots the noise parameters are expressed in terms of the signal-to-noise ratios (SNR), i.e.,

$$\text{SNR}_h = \frac{\|\bar{\mathbf{H}}\|_2^2}{N \frac{1}{\beta}} \quad (46)$$

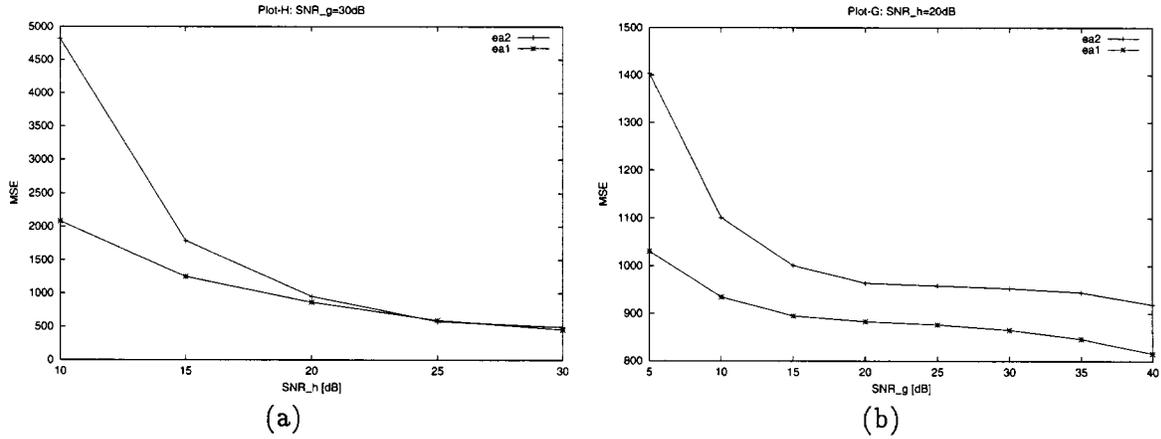


Fig. 1. (a) Experiment I: MSE plot: white PSF noise case, Gaussian random field image, constant- $1/\gamma$. (b) Experiment I: MSE plot: white PSF noise case, Gaussian random field image, constant- $1/\beta$.

where $\|\bar{\mathbf{h}}\|^2$ is the energy of the known part of the PSF, and

$$\text{SNR}_g = \frac{\|\mathbf{f}\|^2}{N \frac{1}{\gamma}} \quad (47)$$

where $\|\mathbf{f}\|^2$ is the energy of the original image.

In all experiments presented in this paper, unless explicitly indicated, the following Gaussian-shaped PSF was used for blurring:

$$h(i, j) = c \cdot \exp\left\{-\frac{i^2 + j^2}{2 \cdot 3^2}\right\}, \quad \text{for } i, j = -15, -14, \dots, -1, 0, 1, \dots, 14, 15 \quad (48)$$

where c is a constant chosen so that $\sum_{i,j} h(i, j) = 1$. The same kernel as in (48) with the additive white-noise component of variance $1/\beta$ was used for restoration. The blurred data was further degraded with additive white observation noise of variance $1/\gamma$. We also performed the experiments where the “smooth” PSF from (48) was used for restoration, while the noisy one ((48) plus the additive PSF noise) was used in the blurring process. Similar results were obtained in both cases.

In all experiments for both the EA1 and the EA2 algorithms after each iteration of the parameter estimation step the functions $L(\alpha, \beta, \gamma)$ in (24) and $\underline{L}(\alpha, \beta, \gamma)$ in (33) are examined, respectively. In all cases we observed that the proposed iterations reduced the value of both functions at each step. The termination of the iterative algorithms was determined based on the convergence of the values of these functions.

Experiment I: In this experiment, we assume white-noise PSF perturbations with exact knowledge of the noise parameters $1/\beta$ and $1/\gamma$ (simulated experiment). The parameter α , however, depends on the original image. In order to have control over α we generate the source image based on the Gaussian image model in (2). More specifically, a white zero-mean Gaussian random field was passed through the linear filter with impulse response equal to $(\alpha \mathbf{Q} \mathbf{Q}^t)^{-\frac{1}{2}}$, where α was preselected and \mathbf{Q} is the circulant Laplacian operator. At the output of this linear

system we obtained the zero-mean Gaussian random field with covariance equal to $(\alpha \mathbf{Q} \mathbf{Q}^t)^{-1}$.

In this experiment, the MSEs for EA1 and EA2 are compared assuming exact knowledge of all three parameters. The constant- $1/\gamma$ and the constant- $1/\beta$ MSE plots are shown in Figs. 1(a) and (b), respectively. From these curves we observe that the EA1 algorithm that uses the exact model outperforms the EA2 algorithm that uses the approximate one. Furthermore, from the constant $1/\gamma$ curve we observe that as the PSF noise gets smaller the performance of the two algorithms becomes similar. This is expected since when the PSF noise becomes zero both algorithms coincide. From the constant $1/\beta$ we observe that as the additive noise gets smaller the difference in performance between the EA1 and EA2 algorithms gets wider. This can be explained by the fact that, in this case, the PSF noise dominates and the EA1 algorithm handles it better.

Experiment II: In this experiment, the “Lena” image was used. In this case, the source image cannot be accurately modeled as a Gaussian random field with the power-spectrum inversely proportional to the transfer function of the Laplacian operator, as is the case with the Gaussian random field in the first experiment. Since exact knowledge of the parameter α is not possible in this case, the proposed algorithms simultaneously estimated the known parameters while restoring the image. More specifically, two cases are shown. In the first, $1/\gamma$ is held constant and is assumed known while α and $1/\beta$ are estimated. In the second, $1/\beta$ is held constant and is assumed known while α and $1/\gamma$ are estimated. The constant- $1/\gamma$ and constant- $1/\beta$ MSE plots are shown in Fig. 2(a) and (b), respectively. From these curves we observe that the EA1 and EA2 algorithms give almost identical performance. The slight difference in performance can be explained by the fact that in this case the SAR image model is also not exact and any difference due to the approximation in $\mathbf{R}_g|_f$ for the EA2 algorithm is not significant.

Experiment III: In this experiment, the EA1 and EA2 approaches are compared under the correlated PSF perturbations. In these experiments we assume that the restoration is performed under PSF modeling errors. More specifically, we assume that the PSFs used for blurring and restoration are Gaussian shaped, but with different widths. The blurring PSF

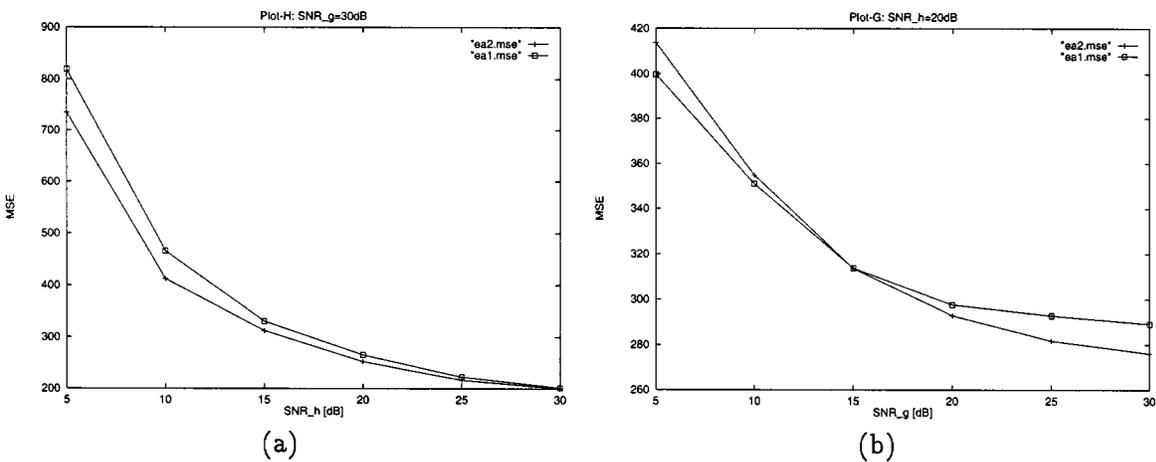


Fig. 2. (a) Experiment II: MSE plot: white PSF noise case, “Lena” image, constant- $1/\gamma$. (b) Experiment II: MSE plot: white PSF noise case, “Lena” image, constant- $1/\beta$.

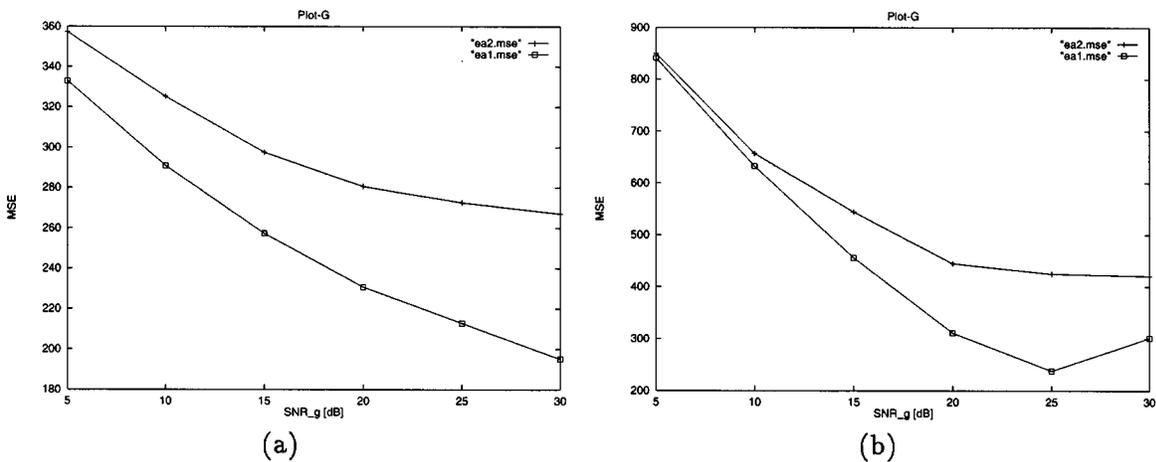


Fig. 3. (a) Experiment III: MSE plot: Correlated PSF noise case, “Lena” image, exact knowledge of power spectrum of the PSF errors. (b) Experiment III: MSE plot: Correlated PSF noise case, “Lena” image, algorithms fit white-noise model to PSF errors.



Fig. 4. Experiment III: Sample images corresponding to the MSE plots in Fig. 3(a): $SNR_g = 25$ dB. (a) Degraded image, (b) EA2 with SAR prior, and (c) EA1 with SAR Prior.

had standard deviation 3.0 while the restoring PSF had standard deviation 4.0. Both PSFs had region of support of 15×15 pixels. The “Lena” image was used in this experiment.

In the first part of this experiment, we compare the MSEs of the EA2 and EA1 restoration filters for SAR image priors assuming exact knowledge of the PSF errors. For this purpose we used the periodogram estimate of the power spectrum of the

PSF errors in both algorithms. The parameter $1/\gamma$ was assumed known for comparison purposes, while the parameter α was estimated simultaneously while restoring. The constant- $1/\gamma$ MSE plot for this experiment is given in Fig. 3(a). Corresponding sample images from this experiment are shown in Fig. 4.

In the second part of this experiment, we assume that no knowledge is available about the PSF errors, and we let the al-

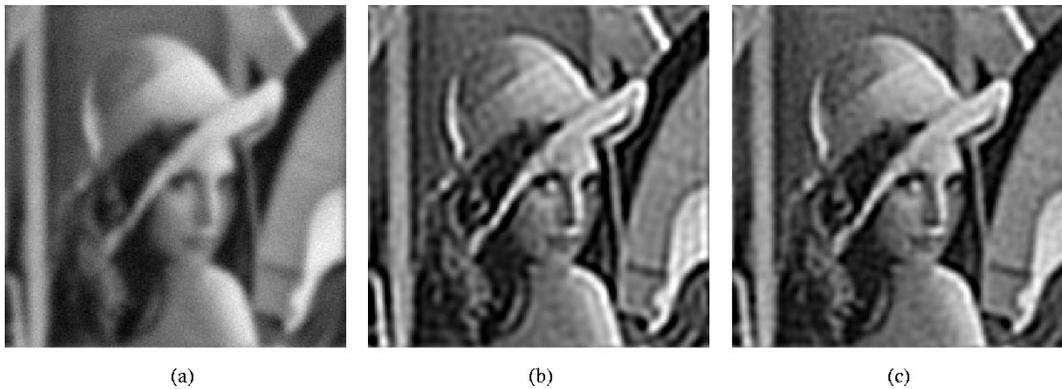


Fig. 5. Experiment III: Sample images Corresponding to the MSE plots in Fig. 3(b): $\text{SNR}_g = 20$ dB; (a) Degraded Image; (b) EA2 with SAR Prior; (c) EA1 with SAR Prior.

gorithms fit a white noise model to the correlated PSF errors. In other words, we simultaneously estimate both parameters α and $1/\beta$ and restore the image. The constant- $1/\gamma$ MSE plot for this experiment is given in Fig. 3(b) and the corresponding images are shown in Fig. 5.

From this experiment we observe that the EA1 algorithm outperforms the EA2 algorithm. This can be explained by the fact that for this experiment where the PSF noise is not white and a white model is used for restoration the approximation introduced in the EA2 algorithm becomes more significant. Furthermore, we observed that the performance gap increases as the additive noise decreases. This behavior was also observed in Experiment I and can be explained similarly.

Experiment IV: In order to compare the proposed approaches with previous ones blind deconvolution based on the EM algorithm [10] using a SAR prior image model was used. This algorithm was compared to the EA2 approach using the same image data as those in Experiment II (the Lena image degraded by the same PSF and with additive white noise). In both algorithms γ was assumed known and the other parameters, α and β for the proposed EA2 approach and α and the PSF for blind deconvolution were estimated. For blind deconvolution the known mean of the PSF was used as the initial value for the PSF estimates. A Monte-Carlo study for $\text{SNR}_g = 20, 30$ and 40 dB with $\text{SNR}_h = 5, 10, 15, 20, 25$ and 30 dB was performed, resulting in 18 different cases total. As before five different noise realizations were used for each case. The constant- γ average MSE plots for $\text{SNR}_g = 40, 30$, and 20 dB are shown in Figs. 6, 7(a), and (b), respectively. From these figures it is clear, as expected, that the EA2 approach that explicitly takes into account the statistics of the PSF model is superior to blind-deconvolution which does not. The interesting observation here is that as the additive noise $\Delta\mathbf{g}$ increases the difference between EM-based blind-deconvolution and EA2 gets more significant. This seems to imply that in EM-based blind-deconvolution the additive noise $\Delta\mathbf{g}$ interferes more with the estimation of the PSF than in the EA2 approach.

In all previous experiments, convergence was achieved in about 50 iterations between the restoration and parameter estimation steps. To implement the restoration step of the EA1 approach we minimized (30) with respect to the real and the imaginary parts of $F(i)$ for every discrete frequency i using

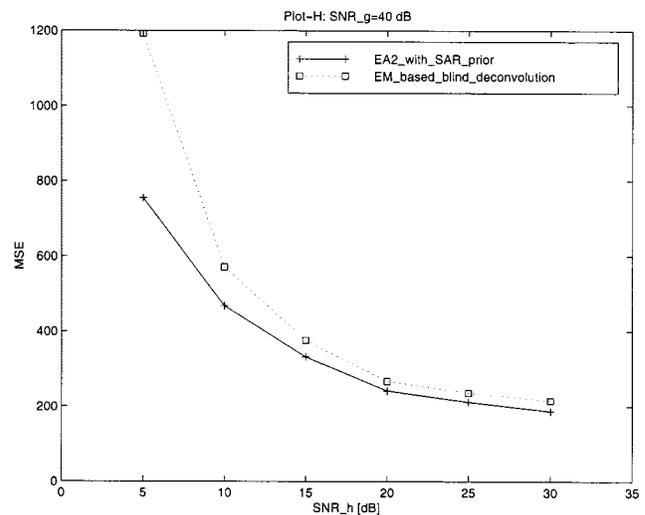


Fig. 6. Experiment IV: Comparison of blind-deconvolution based on the EM with the proposed EA2 approach. The additive noise statistics for both cases are assumed to be known. Constant- γ MSE plot for $\text{SNR}_g = 40$ dB.

the Davidon–Fletcher–Powell optimization algorithm [4]. The gradient required by this algorithm was found in closed form. Convergence of this algorithm was achieved in about ten iterations.

An important observation that was made based on all the numerical experiments that we performed was that neither algorithm could estimate the PSF and the additive noise variances $1/\beta$ and $1/\gamma$ simultaneously. Since the sum of these noise variances appears in the data this was to be expected. To overcome this difficulty prior knowledge in the form of informative priors has to be utilized for α, β, γ .

VIII. CONCLUSIONS AND FUTURE WORK

In this paper, we applied evidence analysis (EA) within the hierarchical Bayesian framework to the parameter estimation and image restoration problem from partially-known blur. In this approach the unknown parameters are treated as random variables (hyper-parameters) with certain probability density functions (PDF), through which the prior knowledge about the hyper-parameters is incorporated into the algorithm. Only the uniform (noninformative) hyper-priors were utilized in this paper. Two

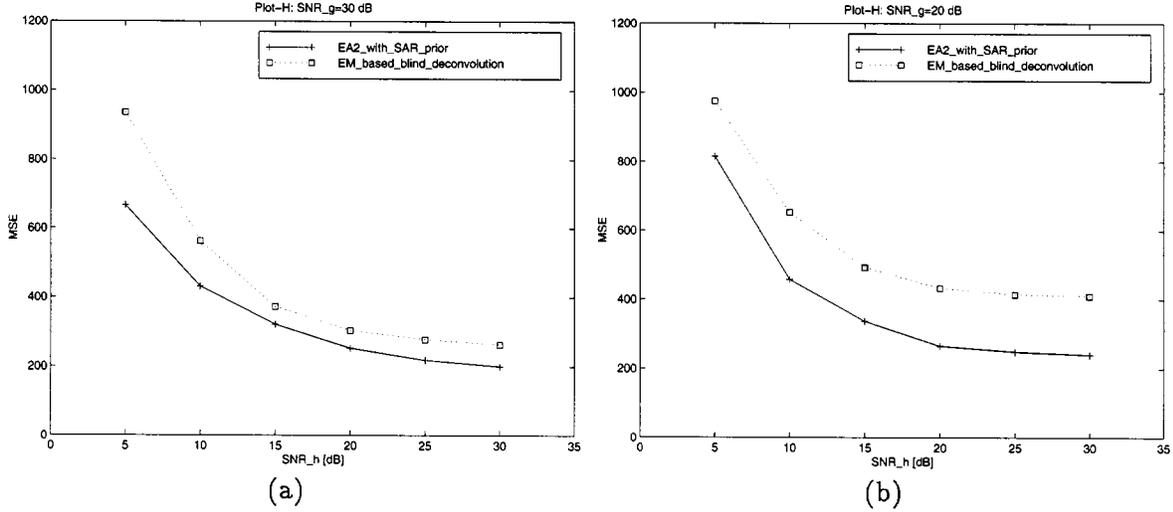


Fig. 7. Experiment IV: Comparison of blind-deconvolution based on the EM with the proposed EA2 approach. The additive noise statistics for both cases are assumed to be known. (a) Constant- γ MSE plot for $\text{SNR}_g = 30$ dB, (b) constant- γ MSE plot for $\text{SNR}_g = 20$ dB.

EA algorithms are derived based on the SAR image model using two different assumptions for $P(\mathbf{g} | \alpha, \beta, \gamma)$. In the restoration step of the first EA algorithm we obtain the image estimate which is the output of the regularized constrained total least squares (RCTLS) filter proposed in [15]. In the restoration step of the second proposed EA algorithm for Gaussian image prior we obtain the image which is the output of the LMMSE filter for this problem. It is interesting to note that when the PSF noise becomes zero both algorithms coincide and become identical to the “iterative” LMMSE filter for the classical restoration problem (PSF completely known) [10], [19]. Therefore, it is expected as the perturbations from the known part of the PSF become smaller the differences in performance between the proposed algorithms and “classical” restoration algorithms diminish.

The two algorithms were experimentally studied under both white and correlated point-spread function (PSF) noise perturbations. Based on the experimental results we conclude that the EA2 algorithm is quite robust to the approximation that the averaged- \mathbf{f} model introduces. Indeed, when the PSF noise is white the EA2 algorithm gives almost identical results with the EA1 algorithm for a range of SNRs. However, when an additional approximation is introduced to the EA2 algorithm as in Experiment III where colored PSF noise is modeled as white noise, then the EA1 algorithm outperforms the EA2 algorithm. Furthermore, the EA1 algorithm worked better in Experiment I where both $\Delta\mathbf{h}$ the PSF noise and \mathbf{f} were Gaussian. This is attributed to the fact that the assumptions used in the EA1 algorithm capture better the statistics of the term $\Delta\mathbf{H}\mathbf{f}$ which was assumed Gaussian for the EA2 algorithm.

Another point that became obvious in our experiments was that neither of the two algorithms can *simultaneously* estimate $1/\beta, 1/\gamma$ when noninformative priors are used. To alleviate this problem, in some initial work on this problem, informative Gamma priors were used (see [16] and [17, ch. 5]). It was then observed that an informative prior for one hyper-parameter only suffices to obtain good estimates for all three of them using the proposed algorithms.

As a final remark, we would like to point out that there is no rigorous proof of convergence for the EA framework. In all our numerical experiments however, we did not experience any convergence difficulties with either one of the proposed algorithms.

APPENDIX A

PARAMETER ESTIMATION STEP FOR THE EA1 ALGORITHM (FIXED- \mathbf{f} COVARIANCE MODEL)

Equation (24) represents the logarithm of the *posterior* PDF which should be minimized with respect to α, β , and γ , according to (10). To perform minimization of (24) we make use of the following matrix algebra formulas:

$$\frac{\partial \mathbf{C}^{-1}(\theta)}{\partial \theta} = -\mathbf{C}^{-1}(\theta) \frac{\partial \mathbf{C}(\theta)}{\partial \theta} \mathbf{C}^{-1}(\theta) \quad (\text{A1})$$

$$\frac{\partial [\log \det \mathbf{C}(\theta)]}{\partial \theta} = \text{tr} \left[\mathbf{C}^{-1}(\theta) \frac{\partial \mathbf{C}(\theta)}{\partial \theta} \right] \quad (\text{A2})$$

where $\mathbf{C}(\theta)$ denotes the square matrix \mathbf{C} which depends on the parameter vector θ , and $\text{tr}(\cdot)$ denotes the trace of a matrix.

Taking the partial derivative of (24) with respect to α , and setting it equal to zero we obtain

$$\frac{N-1}{\alpha} = \left\| \mathbf{Q}\mathbf{f}^{(n)} \right\|^2 + \text{tr} \left[\mathbf{G}^{(n)-1} \mathbf{Q}^t \mathbf{Q} \right]. \quad (\text{A3})$$

Taking the partial derivative of (24) with respect to β we obtain

$$\begin{aligned} & \text{tr} \left[\mathbf{R}_{g|f^{(n)}}^{-1} \frac{1}{\beta^2} \mathbf{F}^{(n)} \mathbf{F}^{(n)t} \right] \\ &= \left(\mathbf{g} - \bar{\mathbf{H}}\mathbf{f}^{(n)} \right)^t \mathbf{R}_{g|f^{(n)}}^{-1} \frac{1}{\beta^2} \mathbf{F}^{(n)} \\ & \quad \times \mathbf{F}^{(n)t} \mathbf{R}_{g|f^{(n)}}^{-1} \left(\mathbf{g} - \bar{\mathbf{H}}\mathbf{f}^{(n)} \right) \\ & \quad + \text{tr} \left[\mathbf{G}^{(n)-1} \bar{\mathbf{H}}^t \mathbf{R}_{g|f^{(n)}}^{-1} \frac{1}{\beta^2} \mathbf{F}^{(n)} \mathbf{F}^{(n)t} \mathbf{R}_{g|f^{(n)}}^{-1} \bar{\mathbf{H}} \right]. \end{aligned} \quad (\text{A4})$$

Using the identity

$$\begin{aligned} \text{tr} \left[\mathbf{R}_{g|f^{(n)}}^{-1} \frac{1}{\beta^2} \mathbf{F}^{(n)} \mathbf{F}^{(n)t} \right] &= \frac{1}{\beta} \left[\text{tr}[\mathbf{I}] - \text{tr} \left[\frac{1}{\gamma} \mathbf{R}_{g|f^{(n)}}^{-1} \right] \right] \\ &= \frac{N}{\beta} - \text{tr} \left[\frac{1}{\beta\gamma} \mathbf{R}_{g|f^{(n)}}^{-1} \right]. \end{aligned} \quad (\text{A5})$$

(A4) can be rewritten as follows:

$$\begin{aligned} \frac{N}{\beta} &= \text{tr} \left[\frac{1}{\beta\gamma} \mathbf{R}_{g|f^{(n)}}^{-1} \right] + \text{tr} \left[\mathbf{G}^{(n)-1} \bar{\mathbf{H}}^t \mathbf{R}_{g|f^{(n)}}^{-1} \right. \\ &\quad \times \left. \frac{1}{\beta^2} \mathbf{F}^{(n)} \mathbf{F}^{(n)t} \mathbf{R}_{g|f^{(n)}}^{-1} \bar{\mathbf{H}} \right] + \left(\mathbf{g} - \bar{\mathbf{H}} \mathbf{f}^{(n)} \right)^t \\ &\quad \times \mathbf{R}_{g|f^{(n)}}^{-1} \frac{1}{\beta^2} \mathbf{F}^{(n)} \mathbf{F}^{(n)t} \mathbf{R}_{g|f^{(n)}}^{-1} \left(\mathbf{g} - \bar{\mathbf{H}} \mathbf{f}^{(n)} \right). \end{aligned} \quad (\text{A6})$$

Similarly, taking the partial derivative of (24) with respect to γ , setting it equal to zero, and using (A5) the following equation is obtained:

$$\begin{aligned} \frac{N}{\gamma} &= \text{tr} \left[\frac{1}{\beta\gamma} \mathbf{F}^{(n)} \mathbf{F}^{(n)t} \mathbf{R}_{g|f^{(n)}}^{-1} \right] \\ &\quad + \text{tr} \left[\mathbf{G}^{(n)-1} \bar{\mathbf{H}}^t \frac{1}{\gamma^2} \mathbf{R}_{g|f^{(n)}}^{-2} \bar{\mathbf{H}} \right] \\ &\quad + \left(\mathbf{g} - \bar{\mathbf{H}} \mathbf{f}^{(n)} \right)^t \frac{1}{\gamma^2} \mathbf{R}_{g|f^{(n)}}^{-2} \left(\mathbf{g} - \bar{\mathbf{H}} \mathbf{f}^{(n)} \right). \end{aligned} \quad (\text{A7})$$

We note that both sides of (A6) are decreasing functions of β for a fixed γ and that both sides of (A7) are decreasing functions of γ for a fixed β . Equations (A3), (A6), and (A7) can then be iteratively solved as shown in (A8)–(A10) at the bottom of the page, where $\mathbf{f}^{(n)}$, $\mathbf{G}^{(n)}$, $\mathbf{F}^{(n)}$, $\alpha^{(n)}$, $\beta^{(n)}$, $\gamma^{(n)}$ are calculated at iteration (n).

APPENDIX B

DERIVATION OF ALGORITHMS IN THE DFT DOMAIN

In this Appendix, we describe the use of the DFT to implement the iterative schemes proposed in this paper. The complete

derivation will be done for the restoration step of the EA1 algorithm which is based on the fixed- \mathbf{f} conditional covariance model; the other implementations follow very similar steps.

Let \mathbf{W} denote the $N \times N$ DFT matrix with $[\mathbf{W}]_{mn} = \exp\{-j(2\pi/N)mn\}$. Denoting by $(\cdot)^{-1}$ and $(\cdot)^H$ the inverse and the Hermitian operations, respectively, $\mathbf{W}^{-1} = (1/N)\mathbf{W}^H$ [12]. Inserting $\mathbf{W}^{-1}\mathbf{W}$ into (29) and using the diagonalization properties of the matrix \mathbf{W} , in other words $\mathbf{W}\mathbf{C}\mathbf{W}^{-1} = \mathbf{D}$, where \mathbf{C} and \mathbf{D} are a circulant matrix and a diagonal matrix with the eigenvalues of \mathbf{C} , we obtain

$$\begin{aligned} J(\mathbf{f}) &= (\bar{\mathbf{H}}\mathbf{f} - \mathbf{g})^H \mathbf{W}^H \frac{1}{N} \frac{1}{\hat{\beta}} \mathbf{W} \\ &\quad \times \left(\mathbf{F}\mathbf{F}^H + \frac{1}{\hat{\gamma}} \mathbf{I} \right)^{-1} \mathbf{W}^{-1} \mathbf{W} (\bar{\mathbf{H}}\mathbf{f} - \mathbf{g}) \\ &\quad + \mathbf{f}^H \mathbf{W}^H \frac{1}{N} \mathbf{W} (\hat{\alpha} \mathbf{Q}\mathbf{Q}^t) \mathbf{W}^{-1} \mathbf{W} \mathbf{f} \\ &\quad + \log \left[\det \left(\mathbf{W} \left(\frac{1}{\hat{\beta}} \mathbf{F}\mathbf{F}^H + \frac{1}{\hat{\gamma}} \mathbf{I} \right) \mathbf{W}^{-1} \right) \right]. \end{aligned} \quad (\text{B1})$$

Now, examining the terms in (B1) separately, and using the diagonalization properties of the DFT for circulant matrices, we obtain

$$\begin{aligned} \mathbf{W}(\bar{\mathbf{H}}\mathbf{f} - \mathbf{g}) &= \mathbf{W}\bar{\mathbf{H}}\mathbf{W}^{-1}\mathbf{W}\mathbf{f} - \mathbf{W}\mathbf{g} = \mathbf{D}_{\bar{\mathbf{H}}}\mathbf{F}_{\text{FT}} - \mathbf{G}_{\text{FT}} \end{aligned} \quad (\text{B2})$$

$$\begin{aligned} \mathbf{W} \left(\frac{1}{\hat{\beta}} \mathbf{F}\mathbf{F}^H + \frac{1}{\hat{\gamma}} \mathbf{I} \right)^{-1} \mathbf{W}^{-1} &= \left[\frac{1}{\hat{\beta}} \mathbf{W}\mathbf{F}\mathbf{W}^{-1}\mathbf{W}\mathbf{F}^H\mathbf{W}^{-1} + \frac{1}{\hat{\gamma}} \mathbf{I} \right]^{-1} = \mathbf{D}_P^{-1} \end{aligned} \quad (\text{B3})$$

$$\det \left(\frac{1}{\hat{\beta}} \mathbf{F}\mathbf{F}^H + \frac{1}{\hat{\gamma}} \mathbf{I} \right) = \prod_{i=0}^{N-1} \left(\frac{1}{\hat{\beta}} |F(i)|^2 + \frac{1}{\hat{\gamma}} \right) \quad (\text{B4})$$

where \mathbf{F}_{FT} and \mathbf{G}_{FT} are the DFT transforms of \mathbf{f} and \mathbf{g} , the eigenvalues of the circulant matrix \mathbf{F} are the DFT values of \mathbf{f} , and $\mathbf{D}_{\bar{\mathbf{H}}}$ and \mathbf{D}_P are given by

$$\mathbf{D}_{\bar{\mathbf{H}}} = \text{diag}[\bar{H}(0), \bar{H}(1), \dots, \bar{H}(N-1)] \quad (\text{B5})$$

$$\frac{1}{\alpha^{(n+1)}} = \left[\frac{\|\mathbf{Q}\mathbf{f}^{(n)}\|^2 + \text{tr} \left[\mathbf{G}^{(n)-1} \mathbf{Q}^t \mathbf{Q} \right]}{N-1} \right] \quad (\text{A8})$$

$$\begin{aligned} \frac{1}{\beta^{(n+1)}} &= \left[\frac{\text{tr} \left[\mathbf{R}_{g|f^{(n)}}^{-1} \right]}{N\beta^{(n)}\gamma^{(n)}} \right] \\ &\quad + \left[\frac{\text{tr} \left[\mathbf{G}^{(n)-1} \bar{\mathbf{H}}^t \mathbf{R}_{g|f^{(n)}}^{-2} \mathbf{F}^{(n)} \mathbf{F}^{(n)t} \bar{\mathbf{H}} \right] + \left(\mathbf{g} - \bar{\mathbf{H}} \mathbf{f}^{(n)} \right)^t \mathbf{R}_{g|f^{(n)}}^{-2} \mathbf{F}^{(n)} \mathbf{F}^{(n)t} \left(\mathbf{g} - \bar{\mathbf{H}} \mathbf{f}^{(n)} \right)}{N\beta^{(n)^2}} \right] \end{aligned} \quad (\text{A9})$$

$$\frac{1}{\gamma^{(n+1)}} = \left[\frac{\text{tr} \left[\mathbf{F}^{(n)} \mathbf{F}^{(n)t} \mathbf{R}_{g|f^{(n)}}^{-1} \right]}{N\beta^{(n)}\gamma^{(n)}} \right] + \left[\frac{\text{tr} \left[\mathbf{G}^{(n)-1} \bar{\mathbf{H}}^t \mathbf{R}_{g|f^{(n)}}^{-2} \bar{\mathbf{H}} \right] + \left(\mathbf{g} - \bar{\mathbf{H}} \mathbf{f}^{(n)} \right)^t \mathbf{R}_{g|f^{(n)}}^{-2} \left(\mathbf{g} - \bar{\mathbf{H}} \mathbf{f}^{(n)} \right)}{N\gamma^{(n)^2}} \right] \quad (\text{A10})$$

and

$$\mathbf{D}_P = \text{diag} \left[\frac{1}{\beta} |F(0)|^2 + \frac{1}{\gamma}, \frac{1}{\beta} |F(1)|^2 + \frac{1}{\gamma}, \dots, \frac{1}{\beta} |F(N-1)|^2 + \frac{1}{\gamma} \right]. \quad (\text{B6})$$

Having established the above, (B1) can be now transformed into

$$J(\mathbf{F}) = \sum_{i=0}^{N-1} \left\{ \frac{1}{N} \left[\frac{|\bar{H}(i)F(i) - G(i)|^2}{\frac{1}{\beta} |F(i)|^2 + \frac{1}{\gamma}} + \hat{\alpha} |Q(i)|^2 |F(i)|^2 \right] + \log \left[\frac{1}{\beta} |F(i)|^2 + \frac{1}{\gamma} \right] \right\}. \quad (\text{B7})$$

Since each term in (B7) is independent of others, the sum in (B7) can be minimized by minimizing each frequency component of the sum separately.

Similar to the derivation of the above restoration step, (25)–(27) can be efficiently implemented in the DFT domain as shown in (B8)–(B10) at the bottom of the page.

APPENDIX C

RELATIONSHIP BETWEEN THE RESTORATION STEPS OF THE TWO EA-BASED ALGORITHMS

We begin by observing that (4) can be rewritten as

$$\mathbf{g} = \bar{\mathbf{g}} + \Delta \tilde{\mathbf{g}} \quad (\text{C1})$$

where

$$\bar{\mathbf{g}} = \bar{\mathbf{H}} \mathbf{f} \quad (\text{C2})$$

and

$$\Delta \tilde{\mathbf{g}} = \Delta \mathbf{H} \mathbf{f} + \Delta \mathbf{g} \quad (\text{C3})$$

is the image-dependent noise term. Now, suppose that the unperturbed (noise-free) system of linear equations $\bar{\mathbf{H}} \mathbf{f} = \bar{\mathbf{g}}$ has a consistent solution \mathbf{f}_0 . Perturb $\bar{\mathbf{g}}$ by $\Delta \tilde{\mathbf{g}}$ to $\mathbf{g} = \bar{\mathbf{g}} + \Delta \tilde{\mathbf{g}}$ around

the consistent solution \mathbf{f}_0 . Then, the consistent solution \mathbf{f}_0 becomes perturbed by $\Delta \mathbf{f}$, where $\Delta \mathbf{f}$ is caused by $\Delta \mathbf{H}$ and $\Delta \mathbf{g}$. The sum of the perturbation error $\Delta \mathbf{f}$ and the consistent solution \mathbf{f}_0 constitute the LMAP solution.

Neglecting the “log” term in (29), the necessary condition for a minimum of (17) is that its gradient is equal to zero. Setting the gradient of $J(\mathbf{f}, \alpha, \beta, \gamma)$ with respect to \mathbf{f} equal to zero yields

$$\begin{aligned} \frac{\partial J(\mathbf{f} | \alpha, \beta, \gamma)}{\partial \mathbf{f}} &= \mathbf{0} = \left[2\bar{\mathbf{H}}^t - (\bar{\mathbf{H}} \mathbf{f} - \mathbf{g})^t \left(\frac{1}{\beta} \mathbf{F} \mathbf{F}^t + \frac{1}{\gamma} \mathbf{I} \right)^{-1} \right. \\ &\quad \left. \times \frac{\partial \left(\frac{1}{\beta} \mathbf{F} \mathbf{F}^t + \frac{1}{\gamma} \mathbf{I} \right)}{\partial \mathbf{f}} \right] \\ &\quad \times \left(\frac{1}{\beta} \mathbf{F} \mathbf{F}^t + \frac{1}{\gamma} \mathbf{I} \right)^{-1} (\bar{\mathbf{H}} \mathbf{f} - \mathbf{g}) + 2\mathbf{R}_f^{-1} \mathbf{f} \end{aligned} \quad (\text{C4})$$

with $\mathbf{R}_f^{-1} = \alpha \mathbf{Q}^t \mathbf{Q}$.

Using $\bar{\mathbf{H}} \mathbf{f}_0 = \bar{\mathbf{g}}$ and neglecting higher-order terms in $\Delta \mathbf{f}$, $\Delta \mathbf{H}$, and $\Delta \mathbf{g}$, and their products, we find from (C4) that

$$\begin{aligned} \frac{\partial J(\mathbf{f})}{\partial \mathbf{f}} = \mathbf{0} &\Leftrightarrow \bar{\mathbf{H}}^t \left(\frac{1}{\beta} \mathbf{F}_0 \mathbf{F}_0^t + \frac{1}{\gamma} \mathbf{I} \right)^{-1} \\ &\quad \times (\bar{\mathbf{H}} \Delta \mathbf{f} - \Delta \mathbf{H} \mathbf{f}_0 - \Delta \mathbf{g}) + \mathbf{R}_f^{-1} (\mathbf{f}_0 + \Delta \mathbf{f}) = \mathbf{0}. \end{aligned} \quad (\text{C5})$$

Letting

$$\mathbf{A} = \left(\bar{\mathbf{H}}^t \left(\frac{1}{\beta} \mathbf{F}_0 \mathbf{F}_0^t + \frac{1}{\gamma} \mathbf{I} \right)^{-1} \bar{\mathbf{H}} + \mathbf{R}_f^{-1} \right)^{-1} \quad (\text{C6})$$

$$\mathbf{B} = \bar{\mathbf{H}}^t \left(\frac{1}{\beta} \mathbf{F}_0 \mathbf{F}_0^t + \frac{1}{\gamma} \mathbf{I} \right)^{-1} \quad (\text{C7})$$

and

$$\mathbf{c} = \mathbf{R}_f^{-1} \mathbf{f}_0 \quad (\text{C8})$$

$$\frac{1}{\alpha^{(n+1)}} = \sum_{i=0}^{N-1} \frac{|Q(i)|^2}{N-1} \left[\frac{\frac{1}{\beta^{(n)}} |F^{(n)}(i)|^2 + \frac{1}{\gamma^{(n)}}}{|\bar{H}(i)|^2 + \alpha^{(n)} \frac{1}{\beta^{(n)}} |Q(i)|^2 |F^{(n)}(i)|^2 + \alpha^{(n)} \frac{1}{\gamma^{(n)}} |Q(i)|^2} + \frac{1}{N} |F^{(n)}(i)|^2 \right] \quad (\text{B8})$$

$$\begin{aligned} \frac{1}{\beta^{(n+1)}} &= \sum_{i=0}^{N-1} \left[\frac{1}{N\beta^2} \frac{|F^{(n)}(i)|^2}{\frac{1}{\beta^{(n)}} |F^{(n)}(i)|^2 + \frac{1}{\gamma^{(n)}}} \frac{|\bar{H}(i)|^2}{|\bar{H}(i)|^2 + \alpha^{(n)} \frac{1}{\beta^{(n)}} |Q(i)|^2 |F^{(n)}(i)|^2 + \alpha^{(n)} \frac{1}{\gamma^{(n)}} |Q(i)|^2} \right. \\ &\quad \left. + \frac{1}{N\beta^{(n)}\gamma^{(n)}} \frac{1}{\frac{1}{\beta^{(n)}} |F^{(n)}(i)|^2 + \frac{1}{\gamma^{(n)}}} + \frac{1}{N\beta^2} \frac{|F^{(n)}(i)|^2}{N} \frac{|G(i) - \bar{H}(i)F^{(n)}(i)|^2}{\left[\frac{1}{\beta^{(n)}} |F^{(n)}(i)|^2 + \frac{1}{\gamma^{(n)}} \right]^2} \right] \end{aligned} \quad (\text{B9})$$

$$\begin{aligned} \frac{1}{\gamma^{(n+1)}} &= \sum_{i=0}^{N-1} \left[\frac{1}{N\gamma^{(n)2}} \frac{1}{\frac{1}{\beta^{(n)}} |F^{(n)}(i)|^2 + \frac{1}{\gamma^{(n)}}} \frac{|\bar{H}(i)|^2}{|\bar{H}(i)|^2 + \alpha^{(n)} \frac{1}{\beta^{(n)}} |Q(i)|^2 |F^{(n)}(i)|^2 + \alpha^{(n)} \frac{1}{\gamma^{(n)}} |Q(i)|^2} \right. \\ &\quad \left. + \frac{1}{N\beta^{(n)}\gamma^{(n)}} \frac{|F^{(n)}(i)|^2}{\frac{1}{\beta^{(n)}} |F^{(n)}(i)|^2 + \frac{1}{\gamma^{(n)}}} + \frac{1}{N\gamma^{(n)2}} \frac{1}{N} \frac{|G(i) - \bar{H}(i)F^{(n)}(i)|^2}{\left[\frac{1}{\beta^{(n)}} |F^{(n)}(i)|^2 + \frac{1}{\gamma^{(n)}} \right]^2} \right] \end{aligned} \quad (\text{B10})$$

and solving (C5) for $\Delta \mathbf{f}$ we obtain

$$\Delta \mathbf{f} = \mathbf{A}[\mathbf{B}(\Delta \mathbf{H} \mathbf{f}_0 + \Delta \mathbf{g}) - \mathbf{c}] \quad (\text{C9})$$

where $(\Delta \mathbf{H} \mathbf{f}_0 + \Delta \mathbf{g})$ is the perturbation $\Delta \tilde{\mathbf{g}}$ around \mathbf{f}_0 . Equation (C9) is a closed form expression for the perturbation from the consistent solution of the estimator obtained from the restoration step of the first EA algorithm. Since it was derived by neglecting higher order terms in $\Delta \mathbf{f}$, $\Delta \mathbf{H}$, $\Delta \mathbf{g}$, and their products, it is valid for small noise levels. We call this the linearized EA1 (LEA1) estimate.

Next, we show that the LEA1 estimate $\hat{\mathbf{f}}_{\text{LEA1}}$ is equal to the estimate $\hat{\mathbf{f}}_{\text{EA2}}$, which is obtained in the restoration step of the second EA algorithm which is also identical to the LMMSE estimate, i.e.,

$$\hat{\mathbf{f}}_{\text{LEA1}} = \mathbf{f}_0 + \Delta \mathbf{f} = \hat{\mathbf{f}}_{\text{EA2}} = \hat{\mathbf{f}}_{\text{LMMSE}}. \quad (\text{C10})$$

Noting that $\hat{\mathbf{f}}_{\text{LEA1}}$ in (C10) can be rewritten as

$$\hat{\mathbf{f}}_{\text{LEA1}} = \mathbf{f}_0 + \mathbf{A}[\mathbf{B}(\mathbf{g} - \bar{\mathbf{H}}\mathbf{f}_0) - \mathbf{c}] \quad (\text{C11})$$

where $\mathbf{g} - \bar{\mathbf{H}}\mathbf{f}_0 = \Delta \mathbf{H} \mathbf{f}_0 + \Delta \mathbf{g} = \Delta \tilde{\mathbf{g}}$, it is straightforward to show using (C6)–(C8) that

$$\hat{\mathbf{f}}_{\text{LEA1}} = \mathbf{A}\mathbf{B}\mathbf{g}. \quad (\text{C12})$$

Recalling (43), we have

$$\hat{\mathbf{f}}_{\text{LMMSE}} = \hat{\mathbf{f}}_{\text{EA2}} = \mathbf{L}\mathbf{g} \quad (\text{C13})$$

where \mathbf{L} is given in (44). Therefore, to show the equivalence of (C12) and (C13) we must show that

$$\mathbf{A}\mathbf{B} = \mathbf{L}. \quad (\text{C14})$$

The verification of (C14) is straightforward using the matrix inversion lemma [11] and the periodogram covariance estimate ($\mathbf{F}_0 \mathbf{F}_0^t = N\mathbf{R}_f$).

REFERENCES

- [1] J. O. Berger, *Statistical Decision Theory and Bayesian Analysis*. New York: Springer-Verlag, 1985.
- [2] P. L. Combettes and H. J. Trussell, "Methods for digital restoration of signals degraded by a stochastic impulse response," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 393–401, Mar. 1989.
- [3] P. L. Combettes, "Signal recovery by best feasible approximation," *IEEE Trans. Image Processing*, vol. 2, pp. 269–271, Apr. 1993.
- [4] B. P. Flannery, W. H. Press, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C*. Cambridge, U.K.: Cambridge Univ. Press, 1988.
- [5] N. P. Galatsanos and A. K. Katsaggelos, "Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation," *IEEE Trans. Image Processing*, vol. 1, pp. 322–336, July 1992.
- [6] 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," *Proc. SPIE*, vol. 3459, pp. 337–348, July 1998.
- [7] L. Guan and R. K. Ward, "Deblurring random time-varying blur," *J. Opt. Soc. Amer. A*, vol. 6, pp. 1727–1737, Nov. 1989.
- [8] —, "Restoration of randomly blurred images by the Wiener filter," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-37, pp. 589–592, Apr. 1989.
- [9] D. Kundur and D. Hatzinakos, "Blind image deconvolution," *IEEE Signal Processing Mag.*, vol. 13, pp. 43–64, May 1996.
- [10] K. T. Lay and A. K. Katsaggelos, "Image identification and restoration based on the expectation-maximization algorithm," *Opt. Eng.*, vol. 29, pp. 436–445, May 1990.
- [11] S. M. Kay, *Fundamentals of Statistical Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [12] A. K. Jain, *Fundamentals of Digital Image Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.
- [13] D. J. C. MacKay, "Comparison of approximate methods for handling hyperparameters," *Neural Computation*, to be published.
- [14] V. Z. Mesarović and N. P. Galatsanos, "MAP and regularized constrained total-least-squares image restoration," in *Proc. IEEE Int. Conf. Image Processing '94*, vol. 3, Austin, TX, Nov. 1994, pp. 177–181.
- [15] V. Z. Mesarović, N. P. Galatsanos, and A. K. Katsaggelos, "Regularized constrained total least-squares image restoration," *IEEE Trans. Image Processing*, vol. 4, pp. 1096–1108, Aug. 1995.
- [16] V. Z. Mesarović, N. P. Galatsanos, and M. N. Wernick, "Restoration from partially-known blur using an expectation-maximization algorithm," in *Proc. 13th ASILOMAR Conf.*, Pacific Grove, CA, Nov. 1996, pp. 95–100.
- [17] V. Z. Mesarović, "Image restoration problems under point-spread function uncertainties," Ph.D. dissertation, ECE Dept., Illinois Inst. Technol., Chicago, May 1997.
- [18] R. Molina, "On the hierarchical Bayesian approach to image restoration: Applications to astronomical images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 16, pp. 1122–1128, Nov. 1994.
- [19] R. Molina, A. K. Katsaggelos, and J. Mateos, "Bayesian and regularization methods for hyperparameters estimation in image restoration," *IEEE Trans. Image Processing*, vol. 8, pp. 231–246, Feb. 1999.
- [20] B. D. Ripley, *Spatial Statistics*. New York: Wiley, 1981.
- [21] J. K. Ruanait and W. J. Fitzgerald, *Numerical Bayesian Methods Applied to Signal Processing*. Berlin, Germany: Springer-Verlag, 1996.
- [22] G. Sharma and H. J. Trussell, "Set theoretic signal restoration using an error in variables criterion," *IEEE Trans. Image Processing*, vol. 6, pp. 1692–1697, Dec. 1997.
- [23] D. Slepian, "Linear least squares filtering of distorted images," *J. Opt. Soc. Amer.*, vol. 57, pp. 918–922, July 1967.
- [24] R. K. Ward and B. E. A. Saleh, "Restoration of images distorted by systems of random impulse response," *J. Opt. Soc. Amer. A*, vol. 2, pp. 1254–1259, Aug. 1985.
- [25] —, "Deblurring random blur," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 35, pp. 1494–1498, Oct. 1987.

Nikolas P. Galatsanos (S'99–M'99) received the Diploma degree in electrical engineering in 1982 from the National Technical University of Athens, Athens, Greece, the M.S.E.E. and Ph.D. 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 presently he holds the rank of Associate Professor and serves as the Director of the graduate program. His current research interests center around image/signal processing and computational vision problems for visual communications and medical imaging applications. He is co-editor of *Recovery Techniques for image and Video Compression and Transmission* (Norwell, MA: Kluwer, 1998).

Dr. Galatsanos served as Associate Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING from 1992 to 1994. He is an Associate Editor for the *IEEE Signal Processing Magazine*.



Vladimir Z. Mesarović was born in Belgrade, Yugoslavia, in 1967. He received the Diploma degree in electrical engineering from the University of Belgrade in 1992, and the M.S. and the Ph.D. 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/decoding, transmission and their efficient DSP implementations.



Rafael Molina was born in 1957. He received the degree in mathematics (statistics) in 1979 and the Ph.D. degree in in optimal design in linear models in 1983.

He became Associate Professor of computer science and artificial intelligence at the University of Granada, Granada, Spain, in 1989, 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, Royal Statistical Society, and the Asociación Española de Reconocimiento de Formas y Análisis de Imágenes (AERFAI).



Aggelos K. Katsaggelos (S'80–M'85–SM'92–F'98) received the Diploma degree in electrical and mechanical engineering from the Aristotelian University of Thessaloniki, Thessaloniki, Greece, in 1979 and the M.S. and Ph.D. degrees both in electrical engineering from the Georgia Institute of Technology, Atlanta, GA, in 1981 and 1985, respectively.

In 1985, he joined the Department of Electrical Engineering and Computer Science, Northwestern University, Evanston, IL, where he is currently 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 with the Department of Electrical Engineering and Computer Science, Polytechnic University, 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. He was an Area Editor for the *Graphical Models and Image Processing* (1992–1995). He is the editor of *Digital Image Restoration* (Berlin, Germany: Springer-Verlag, 1991), co-author of *Rate-Distortion Based Video Compression* (Norwell, MA: Kluwer, 1997), and co-editor of *Recovery Techniques for Image and Video Compression and Transmission*, (Norwell, MA: Kluwer, 1998). He served as the General Chairman of the 1994 Visual Communications and Image Processing Conference, Chicago, IL. He is the co-inventor of seven international patents with several pending.

Dr. Katsaggelos is an Ameritech Fellow, a member of the Associate Staff, Department of Medicine, Evanston Hospital, and a member of SPIE. He is a member of the Board of Governors and the Publication Board of the IEEE Signal Processing Society, the IEEE TAB Magazine Committee, the IEEE Technical Committee on Visual Signal Processing and Communications, and Multimedia Signal Processing, and Editor-in-Chief of the *IEEE Signal Processing Magazine*. He served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING (1990–1992), a member of the Steering Committees of the IEEE TRANSACTIONS ON IMAGE PROCESSING (1992–1997) and the IEEE TRANSACTIONS ON MEDICAL IMAGING (1990–1999), and a member of the IEEE Technical Committee on Image and Multi-Dimensional Signal Processing (1992–1998). He was technical program co-chair of the 1998 IEEE International Conference on Image Processing, Chicago, IL. He is the recipient of the IEEE Third Millennium Medal.