HIERARCHICAL BAYESIAN IMAGE RESTORATION FROM PARTIALLY-KNOWN BLURS

Vladimir Mesarović¹, Nikolas Galatsanos², Rafael Molina³, Aggelos Katsaggelos⁴

¹Crystal Semiconductor Corporation, Austin, TX 78741
 ²Illinois Institute of Technology, Chicago, IL 60616
 ³Universitad de Granada, 18071 Granada, Espana
 ⁴Northwestern University, Evanston, IL 60208

ABSTRACT

A number of restoration filters have been proposed for the restoration problem from partially-known blurs. Recently we proposed the regularized constrained least-squares filter (RCTLS) and we showed that it has a number of advantages over previous ones [4]. However, the problem of estimating the parameters that define the RCTLS filter has not yet been addressed. In this paper we propose a two-step algorithm based on the hierarchical Bayesian approach to simultaneously restore the image and estimate the parameters of the RCTLS restoration filter. The algorithm is derived in the DFT domain; thus, it is very efficient even for very large images.

1. INTRODUCTION

In most practical applications, the point-spread function (PSF) is neither unknown nor perfectly known. For instance, in medical imaging techniques such as positron emission tomography (PET) and single-photon emission computed tomography (SPECT), the PSF is difficult to specify completely, in part because it is object-dependent, owing to scattering and photon attenuation. In astronomy, atmospheric turbulence yields a random time-varying PSF which is not known exactly when the image is restored.

Random blurs have been considered before. A linear minimum mean square error filter (LMMSE) is developed in [3, 8]. In [8] an "ad-hoc" iterative algorithm for covariance estimation is proposed. The convergence properties of this algorithm, however, were not analyzed in [8].

In [4] fast regularized constrained total least squares (RCTLS) filters were derived and implemented in the discrete Fourier transform (DFT) domain. It was shown in [4] that the RCTLS filters were superior to the regularized-least-squares (RLS) filters that do not take into account the errors in the PSF and also to the linear minimum mean square error (LMMSE) filter [8] for this problem. However, to effectively utilize the RCTLS filter in [4], the noise and the image prior parameters must be known prior to restoration.

In [5, 6] two iterative schemes for simultaneous parameter estimation and image restoration based on the Expectation Maximization (EM) algorithm were proposed. Although the algorithms in [5, 6] were shown to be very powerfull, they were derived under the assumption that the observed data on the image-dependent noise term are Gaussian which may be too restrictive in certain cases. Because of this assumption the restoration step of that algorithm yields a linear filter which is identical to the LMMSE filter in [8].

To overcome the above difficulties for this problem, in this paper we apply Evidence Analysis (EA), within the hierarchical Bayesian framework [2, 7], to the partiallyknown blur restoration problem. The EA yields the RCTLS restoration filter in [4] and simultaneously estimates the required parameters.

The rest of this paper is organized as follows: In section II the observation and the image models are discussed. In section III the hierarchical Bayesian analysis is introduced and the EA algorithm is derived. Numerical experiments are presented in section IV and conclusions are given in section V.

2. OBSERVATION AND IMAGE MODELS

We assume the following degradation model [3, 4, 8]:

$$\mathbf{g} = \mathbf{H} \, \mathbf{f} + \mathbf{\Delta} \mathbf{g},\tag{1}$$

in which $\mathbf{g}, \mathbf{f}, \Delta \mathbf{g} \in \mathcal{R}^N$ are lexicographically ordered representations of the observed degraded image, the source image, and the additive noise in the observed image, respectively. The space-invariant PSF **H** is represented as the sum of a deterministic component $\mathbf{\bar{H}}$ and a stochastic component of zero-mean $\Delta \mathbf{H}$, i.e.,

$$\mathbf{H} = \mathbf{\bar{H}} + \mathbf{\Delta}\mathbf{H}.$$
 (2)

The matrix $\tilde{\mathbf{H}}$ 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. The unknown component of the PSF is modeled as stationary zero-mean white noise with $N \times N$ covariance matrix $\mathbf{R}_{\Delta h} = \frac{1}{\beta}\mathbf{I}$, where $\frac{1}{\beta}$ denotes the variance of the PSF noise and \mathbf{I} is the identity matrix. The observation vector \mathbf{g} is also modeled with $N \times N$ covariance matrix $\mathbf{R}_{\Delta g} = \frac{1}{\gamma}\mathbf{I}$, where $\frac{1}{\gamma}$ denotes the variance of the observation vector \mathbf{g} is also modeled with $N \times N$ covariance matrix $\mathbf{R}_{\Delta g} = \frac{1}{\gamma}\mathbf{I}$, where $\frac{1}{\gamma}$ denotes the variance of the observation noise. A circulant approximation of Toeplitz matrices [1] will be used to allow calculations to be performed using the discrete Fourier transform (DFT); thus, $\mathbf{R}_{\Delta h}$, $\mathbf{R}_{\Delta g}$, $\mathbf{\tilde{H}}$ and $\Delta \mathbf{H}$ are $N \times N$ circulant matrices [1].

To reduce the complexity of the parameter estimation step, we use the simultaneously autoregressive (SAR) image prior model [7]. This model can be described by the following conditional PDF:

$$P(\mathbf{f}|\alpha) = \operatorname{const} \cdot \alpha^{(\frac{N}{2})} \exp\left\{-\frac{\alpha}{2}||\mathbf{Q}\mathbf{f}||^2\right\},\qquad(3)$$

where α is positive unknown parameter that controls the smoothness of the image and $||\mathbf{Q}\mathbf{f}||^2$ is a non negative quadratic form that captures the image autoregressive model. 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 [1]. Any other quadratic image prior can be used without loss of generality (see [6] for Gaussian priors).

3. HIERARCHICAL BAYESIAN ANALYSIS

For the problem at hand, it is easy to see that the *posterior* distribution for the image f depends on the unknown parameters that describe the observation and the image models. In the hierarchical Bayesian framework those parameters can be treated as random variables (hyperparameters), and as such they are assigned hyperpriors (also called a multistage priors) [2]. We will assume uniform (noninformative) hyperpriors on the unknown parameters, defined on $[0, \infty)$.

According to the Evidense Analysis (EA) approach, in the hierarchical Bayesian framework [7], the simultaneous estimation of \mathbf{f} , α , β , and γ can be done in two steps: Parameter estimation step:

$$\hat{\alpha}, \hat{\beta}, \hat{\gamma} = \arg \max_{\alpha, \beta, \gamma} \left\{ P(\alpha, \beta, \gamma | \mathbf{g}) \right\}.$$
 (4)

Restoration step:

$$\hat{\mathbf{f}}(\hat{\alpha},\hat{\beta},\hat{\gamma}) = \arg \max_{\mathbf{f}} \left\{ P(\mathbf{f}|\mathbf{g},\hat{\alpha},\hat{\beta},\hat{\gamma}) \right\}.$$
(5)

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

According to Bayes rule for PDFs, the joint PDF of the data, the image, and the parameters can be represented as

$$P(\mathbf{g}, \mathbf{f}, \alpha, \beta, \gamma) = P(\mathbf{g} | \mathbf{f}, \alpha, \beta, \gamma) P(\mathbf{f} | \alpha, \beta, \gamma) P(\alpha) P(\beta) P(\gamma),$$
(6)

where the hyperparameters are assumed independent from each other, $P(\mathbf{f}|\alpha,\beta,\gamma)$ does not depend on β and γ and is given in (3). Then, to obtain $P(\alpha,\beta,\gamma|\mathbf{g})$, as required by (4), we marginalize the PDF in (6) with respect to \mathbf{f} [2, 7], i.e.,

$$P(\alpha,\beta,\gamma|\mathbf{g}) \propto \int P(\mathbf{g},\mathbf{f},\alpha,\beta,\gamma) d\mathbf{f}.$$
 (7)

Since we assumed "flat" (non-informative) hyperpriors, $P(\alpha)P(\beta)P(\gamma)$ can be discarded in (6). For the restoration step, as required in (5), the image posterior $P(\mathbf{f}|\mathbf{g}, \hat{\alpha}, \hat{\beta}, \hat{\gamma})$ can be obtained applying the Bayes rule to the joint PDF, i.e.,

$$P(\mathbf{f}|\mathbf{g}, \hat{\alpha}, \hat{\beta}, \hat{\gamma})P(\mathbf{g}, \hat{\alpha}, \hat{\beta}, \hat{\gamma}) = P(\mathbf{g}|\mathbf{f}, \hat{\alpha}, \hat{\beta}, \hat{\gamma})P(\mathbf{f}|\hat{\alpha}, \hat{\beta}, \hat{\gamma}),$$
(8)

where $P(\mathbf{g}, \hat{\alpha}, \beta, \hat{\gamma})$ does not depend on \mathbf{f} , and $P(\mathbf{f}|\hat{\alpha}, \beta, \hat{\gamma})$ is given by (3), evaluated at $\hat{\alpha}, \hat{\beta}$, and $\hat{\gamma}$.

We assume Gaussian distribution for the *noise* both in the PSF and the observations. Then, to determine $P(\mathbf{g}|\mathbf{f}, \alpha, \beta, \gamma)$ we note that vector \mathbf{f} is not a random quantity, but rather a fixed one. It is easy to see from (1) that the likelihood $P(\mathbf{g}|\mathbf{f}, \alpha, \beta, \gamma)$ is a Gaussian PDF with mean equal to $\mathbf{\bar{H}}\mathbf{f}$, i.e.,

$$P(\mathbf{g}|\mathbf{f},\alpha,\beta,\gamma) = \left[\det(2\pi\mathbf{R}_{g|f})\right]^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{g}-\mathbf{H}\mathbf{f})^{t}\mathbf{R}_{g|f}^{-1}(\mathbf{g}-\mathbf{H}\mathbf{f})\right\}.$$
(9)

The conditional covariance $\mathbf{R}_{g|f}$ in (9) is given by

$$\mathbf{R}_{g|f} = E \left\{ (\mathbf{\Delta}\mathbf{H}\mathbf{f} + \mathbf{\Delta}\mathbf{g})(\mathbf{\Delta}\mathbf{H}\mathbf{f} + \mathbf{\Delta}\mathbf{g})^{t} \right\} \\ = E \left\{ (\mathbf{F}\mathbf{\Delta}\mathbf{h} + \mathbf{\Delta}\mathbf{g})(\mathbf{F}\mathbf{\Delta}\mathbf{h} + \mathbf{\Delta}\mathbf{g})^{t} \right\},$$
(10)

where we have used the commutative property of the convolution operation, \mathbf{F} denotes the circulant matrix generated by the image \mathbf{f} , and $\Delta \mathbf{h}$ is the PSF noise vector that generates $\Delta \mathbf{H}$. Equation (10) can be further simplified as follows:

$$\mathbf{R}_{g|f} = \mathbf{F}E\left\{\mathbf{\Delta}\mathbf{h}\mathbf{\Delta}\mathbf{h}^{t}\right\}\mathbf{F}^{t} + E\left\{\mathbf{\Delta}\mathbf{g}\mathbf{\Delta}\mathbf{g}^{t}\right\} = \frac{1}{\beta}\mathbf{F}\mathbf{F}^{t} + \frac{1}{\gamma}\mathbf{I},$$
(11)

where $\frac{1}{\beta}$ and $\frac{1}{\gamma}$ are the PSF- and the observation noise variances.

Substituting (3) and (9) into (6) we obtain

$$P(\mathbf{g}, \mathbf{f}, \alpha, \beta, \gamma) \propto \alpha^{(\frac{N}{2})} [\det(\mathbf{R}_{g|f})]^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}J(\mathbf{f}, \alpha, \beta, \gamma)\right\}$$
(12)

where

$$J(\mathbf{f},\alpha,\beta,\gamma) = \alpha ||\mathbf{Q}\mathbf{f}||^2 + (\mathbf{g} - \mathbf{\bar{H}}\mathbf{f})^{t}\mathbf{R}_{g|f}^{-1}(\mathbf{g} - \mathbf{\bar{H}}\mathbf{f}). \quad (13)$$

3.1. Restoration Step

According to Bayes law

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

Since $P(\mathbf{g}, \alpha, \beta, \gamma)$ does not depend on **f**

$$\arg\min_{\mathbf{f}} \{P(\mathbf{f}|\mathbf{g},\alpha,\beta,\gamma)\} = \arg\min_{\mathbf{f}} \{P(\mathbf{f},\mathbf{g},\alpha,\beta,\gamma)\}.$$
(15)

Substituting (12) and (13) into (15) we get

$$\hat{\mathbf{f}}(\hat{\alpha}, \hat{\beta}, \hat{\gamma}) = \arg\min_{\mathbf{f}} \left(\bar{\mathbf{H}} \mathbf{f} - \mathbf{g} \right)^{t} \hat{\mathbf{R}}_{g|f}^{-1} (\bar{\mathbf{H}} \mathbf{f} - \mathbf{g}) + \hat{\alpha} ||\mathbf{Q} \mathbf{f}||^{2} + \log \left[\det(\hat{\mathbf{R}}_{g|f}) \right],$$
(16)

where $\hat{\mathbf{R}}_{g|f} = \frac{1}{\hat{\beta}} \mathbf{F} \mathbf{F}^t + \frac{1}{\hat{\gamma}}$.

Equation (16) is equivalent to the RCTLS restoration filter in [4]. A practical computation of (16) can be obtained

by transformation to the DFT domain and ignoring the log term in (16) [6] yields

$$\begin{split} \hat{F}(i) &= \arg\min_{F(i)} \frac{1}{N} \left[\frac{|\tilde{H}(i)F(i)-G(i)|^2}{\frac{1}{\beta}|F(i)|^2 + \frac{1}{\gamma}} + \hat{\alpha} |Q(i)|^2 |F(i)|^2 \right] \\ &+ \log \left[\frac{1}{\beta} |F(i)|^2 + \frac{1}{\gamma} \right], \end{split}$$
(17)

for each frequency $i = 0, 1, \dots, N-1$. In (17) G(i) and F(i) are the DFT coefficients of the observed and the restored images, $\bar{H}(i)$ and Q(i) are the eigenvalues of $\bar{\mathbf{H}}$ and \mathbf{Q} , and $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\gamma}$ are the estimates of the hyperparameters obtained in the parameter estimation step.

3.2. Parameter Estimation Step

To compute $P(\alpha, \beta, \gamma | \mathbf{g})$ we substitute (12) into (7). Then, expanding $J(\mathbf{f}, \alpha, \beta, \gamma)$ in Taylor series (second order) around a known $\mathbf{f}^{(n)}$ ((n) denotes the iteration index), and performing the integration term by term we obtain:

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

where

$$\mathbf{G}^{(n)} = \alpha \mathbf{Q}^{t} \mathbf{Q} + \mathbf{\hat{H}}^{t} \mathbf{R}_{g|f^{(n)}}^{-1} \mathbf{\hat{H}}.$$
 (19)

Taking "2 log" of both sides of (18) and taking partial derivatives of the obtained posterior functional with respect to α , β , and γ , we obtain:

$$\frac{N}{\alpha} = ||\mathbf{Q} \mathbf{f}^{(n)}||^2 + \operatorname{tr}[\mathbf{G}^{(n)^{-1}}\mathbf{Q}^{t}\mathbf{Q}]$$
(20)

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

$$(21)$$

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

$$+ (\mathbf{g} - \mathbf{\bar{H}}\mathbf{f}^{(n)})^{t}\frac{1}{\gamma^{2}}\mathbf{R}_{g|f(n)}^{-2}(\mathbf{g} - \mathbf{\bar{H}}\mathbf{f}^{(n)}).$$

$$(22)$$

The parameter estimation cycle in (20)-(22) is repeated in the DFT domain since all matrices are circulant, until convergence in (18) occurs. In all our experiments with the EA algorithm we observed that the convergence occured after several iterations.

4. NUMERICAL EXPERIMENTS

In this section we test the proposed EA algorithm and compare it to the EM algorithm in [5].

The (per pixel) MSE is defined as $MSE = \frac{1}{N} \|\mathbf{f} - \mathbf{f}\|_2^2$, where \mathbf{f} and $\mathbf{\hat{f}}$ are the original and the restored (upon convergence) images, respectively. The MSE measurements were performed based on Monte-Carlo simulations. To avoid plotting the 3-D plot of the MSE versus both noise parameters we plot two 2-D MSE plots: (Plot-H): For a fixed $SNR_g = 30dB$ we plot MSE versus SNR_h by varying $\frac{1}{\beta}$, and (Plot-G): For a fixed $SNR_h = 20dB$ we plot MSE versus SNR_g by varying $\frac{1}{\gamma}$. In those plots the noise parameters are expressed in terms of the signal-to-noise ratios (SNR), i.e.,

$$\mathrm{SNR}_{h} = \frac{\|\mathbf{\bar{h}}\|^{2}}{N\frac{1}{\beta}}; \ \mathrm{SNR}_{g} = \frac{\|\mathbf{f}\|^{2}}{N\frac{1}{\gamma}}, \tag{23}$$

where $\|\bar{\mathbf{h}}\|^2$ and $\|\mathbf{f}\|^2$ are the energies of the known part of the PSF, and the original image, respectively.

In all experiments presented in this paper Gaussianshaped PSF given below was used for blurring:

$$h(i,j) = c \cdot \exp\left\{-rac{i^2+j^2}{2\cdot 3^2}
ight\}$$
, for $i,j = -15,\cdots,15,$

where c is a constant chosen so that $\sum_{i,j} h(i,j) = 1$.

Experiment I. In this experiment we assume whitenoise PSF perturbations with exact knowledge of the noise parameters $\frac{1}{\beta}$ and $\frac{1}{\gamma}$. In order to have a control over parameter α we generated the source image based on the Gaussian image model in (3) and preselected value for α . Then, assuming the knowledge of all three parameters, the MSEs of the EA and the EM restoration filters are calculated in Figures 1 and 2.

Experiment II. In this experiment EA and the EM approaches are compared under the correlated PSF perturbations. More specifically, we assume that the PSFs used for blurring and restoration are Gaussian shaped, but with different widths. The blurring PSF had standard deviation 3.0 while the restoring PSF had standard deviation 4.0. "Lena" image was used in this experiment. The exact knowledge of the spectrum of the PSF errors and the knowledge of the parameter $\frac{1}{\gamma}$ were assumed, while parameter α was estimated simultaneously while restoring. The constant- $\frac{1}{\gamma}$ MSE plot is given in Figure 3 and the corresponding images are shown in Figure 4.

Based on the performed experiments we make the following conclusions. (1) If the source image can be accurately modeled with the prior in (3), the EM-based restoration algorithm (linear) is inferior to the EA restoration filter (nonlinear), since the Gaussian assumption on the imagedependent term does not capture the statistics of the degradation model. (2) The total noise term in the observation model for the correlated noise case cannot be accurately modeled with the Gaussian PDF. The images obtained with the EM approach in this case have ringing artifacts and poor resolution as compared to images obtained with the EA approach. (3) In all our experiments neither algorithm could estimate the PSF and additive noise variances $\frac{1}{\beta}$ and $\frac{1}{\gamma}$ simultaneously, since the sum of these noise parameters appears in the data. Gamma priors were successfully introduced in [6] to alleviate this problem.

5. CONCLUSIONS

In this paper we applied Empirical Bayesian (EA) analysis to the simultaneous parameter estimation and image restoration problem from partially-known blurs. The proposed algorithm was experimentally demonstrated under both white and correlated point-spread function (PSF) noise perturbations.

6. REFERENCES

- [1] Andrews H. and Hunt B., *Digital Image Restoration*, Englewood Cliffs, NJ: Prentice-Hall, 1977.
- Berger J. O., Statistical Decision Theory and Bayesian Analysis, Springer-Verlag, New York, 1985.
- [3] Guan L. and Ward R. K., "Deblurring Random Time-Varying Blur", Journal of the Optical Society of America A., Vol. 6, No. 11, pp. 1727-1737, November 1989.
- [4] Mesarović V. Z., Galatsanos N. P., and Katsaggelos A. K., "Regularized Constrained Total Least-Squares Image Restoration", *IEEE Trans. Image Processing*, Vol. 4, No. 8, pp. 1096-1108, August 1995.
- [5] Mesarović V. Z., Galatsanos N. P., and Wernick M. N., "Restoration from Partially-Known Blur Using an Expectation-Maximization Algorithm", *Proceedings of Thirtieth ASILOMAR conference*, pp. 95-100, Pacific Grove, November 1996.
- [6] Mesarović V. Z., Image Restoration Problems Under Point-Spread Function Uncertainties, Ph. D. Thesis, ECE Dept., Illinois Institute of Technology, Chicago, May 1997.
- [7] Molina R., "On the Hierarchical Bayesian Approach to Image Restoration: Applications to Astronomical Images", *IEEE Trans. on Pattern Analysis and Machine Inteligence*, Vol. 16, No. 11, pp. 1122-1128, November 1994.
- [8] Ward R. K. and Saleh B. E. A., "Restoration of Images Distorted by Systems of Random Impulse Response", *Journal of the Optical Society of America A.*, Vol. 2, No.8, pp. 1254-1259, August 1985.





Figure 3. Correlated PSF noise



Figure 4(a). Degraded image



Figure 4(b). EM algorithm



Figure 4(c) EA algorithm