

# Fast Rauch–Tung–Striebel Smoother-Based Image Restoration for Noncausal Images

Amir Asif, *Senior Member, IEEE*

**Abstract**—We describe a technique for restoration of blurred images corrupted with additive noise. Our algorithm uses a practical implementation of the Rauch–Tung–Striebel (RTS) smoother-based on noncausal prediction that models the blurred image as a finite-lattice Gauss Markov random process (GMRP). The one-sided regressors of the GMRP converge at a geometric rate to shift-invariant values along the rows of the image. This leads to a steady-state solution for the RTS filter. Experimental results illustrate the superiority of our RTS-based algorithm over Wiener filter, deterministic filters, and filters that use the one-sided causal state model.

**Index Terms**—Gauss–Markov random process (GMRP), image restoration, Kalman–Bucy filter (KBF), noncausal prediction, random process, Rauch–Tung–Striebel (RTS) smoother.

## I. INTRODUCTION

**A**N IMAGE ACQUIRED by a practical (typically imperfect) imaging system suffers from degradations resulting from such factors as sensor noise, improper camera focus, and relative object–camera motion. If the imaging system is modeled by a linear shift-invariant (LSI) system, the observed image  $z(i, j)$  is expressed in terms of the ideal image  $x(i, j)$  as

$$z(i, j) = g(i, j) \oplus x(i, j) + v(i, j) \quad (1)$$

where  $\oplus$  denotes the two-dimensional (2-D) convolution operator,  $g(i, j)$  is the point spread function (psf) that introduces blur and distortion in  $x(i, j)$ , and  $v(i, j)$  represents additive noise. The objective of image restoration is to determine  $x(i, j)$  from the observed image  $z(i, j)$ .

In recent years, a number of robust algorithms have been proposed for image restoration. Broadly speaking these can be categorized in two categories. Category 1 based on inverse filters, [2], considers only blurring and performs poorly in the presence of the observation noise. Category 2 proposes the Wiener filter or compound Gauss Markov random field (CGMRF)-based restoration techniques [3], to control changes in the image model using a hidden random field. The process of finding the maximum *a posteriori* (MAP) estimate is complex and methods like simulated annealing, deterministic relaxation, or anisotropic diffusion are used. Such methods are computationally intensive and lead to unstable solutions.

The letter describes a fast implementation of the Rauch–Tung–Striebel (RTS) smoother [1] for image restoration

Manuscript received February 19, 2003; revised June 25, 2003. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Stanley J. Reeves.

The author is with the Department of Computer Science, York University, North York, ON, Canada M3J 1P3 (e-mail: asif@cs.yorku.ca).

Digital Object Identifier 10.1109/LSP.2003.822922

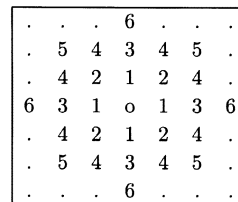


Fig. 1. First- to sixth-order neighborhood noncausal prediction models.

based on a noncausal predictive model. Our approach models the image  $x(i, j)$  as a noncausal finite-lattice Gauss Markov random process (GMRP), [4]–[7]. We extend the work presented in [6] and restore a noisy image blurred with a known psf using a computationally efficient implementation of the RTS algorithm. Our work supplements [6] in two ways. While [6] is limited to images corrupted by additive noise (no blur is considered), we consider degradations introduced from both noise and blur. A second contribution of this work is to derive a practical implementation of the RTS algorithm. We make use of the fact that the state matrices in the state model obtained from the noncausal GMRP predictive model reach shift-invariant values along the rows of the noisy image. The predictor and filter covariance matrices computed in the forward stage of the Kalman–Bucy filter (KBF) therefore attain steady-state values after a few iterations. Once this happens, the smoother gain and smoother covariance in the backward sweep of the RTS algorithm do not require any further updating. This reduces the computational complexity of the RTS algorithm.

The letter is organized as follows. Section II reviews the noncausal GMRP used to model the image. Under mild constraints, the regressors used to define the state model of the noncausal image field reach shift-invariant values. Section III describes the psfs that we consider. In Section IV, the RTS algorithm is presented with a note on its steady-state values. Section V compares the performance of the algorithm with spatial averaging, Wiener filter, and a similar RTS algorithm based on causal prediction. Finally, Section VI concludes the letter.

## II. NONCAUSAL GMRP

In noncausal GMRP, we use a noncausal neighborhood of pixels to make a linear prediction of the current pixel value. Fig. 1 illustrates the noncausal neighborhood sets of order up to six for the pixel labeled o. The first-order GMRP involves pixels marked 1, the second-order set involves the pixels marked as 1 and 2, and similarly for the higher order processes. For the first-order GMRP, the prediction of the image intensity at pixel

$(i, j)$  (say the pixel represented by  $o$  in Fig. 1) given the image intensity values at all other pixels, is

$$\hat{x}(i, j) = \beta_v x(i-1, j) + \beta_h x(i, j-1) + \beta_h x(i, j+1) + \beta_v x(i+1, j) \quad (2)$$

where  $\beta_v$  and  $\beta_h$  are the vertical and horizontal field interactions,  $x(i-1, j)$  represents the intensity value of the pixel in the row above the reference pixel, row  $(i-1)$ , and in the same column as the reference pixel, column  $j$ , of an  $(I \times J)$  image and likewise for the remaining quantities in (2). The error  $e(i, j)$  involved in the prediction is given by  $e(i, j) = x(i, j) - \hat{x}(i, j)$ . Using the row-major order to vectorize image  $X_i$  by vector  $\vec{X}$  and error field  $e(i, j)$  by  $\vec{e}$ , the error image can be expressed in the matrix-vector form,  $\mathcal{A}\vec{X} = \vec{e}$ , where  $\vec{e}$  is the  $(IJ \times 1)$  column vector for the error image  $e(i, j)$ . The matrix  $\mathcal{A}$  is the  $(IJ \times IJ)$  potential matrix and has the following structure for zero Dirichlet<sup>1</sup> boundary conditions (BCs)

$$\mathcal{A} = I_I \otimes B + H_I^1 \otimes C \quad \text{with } B = -\beta_h H_J^1 + I_J \quad \text{and } C = -\beta_v I_J \quad (3)$$

where  $\otimes$  represent the Kronecker product. The symbols  $I_I$  and  $I_J$  are identity matrices, while  $H_I^1$  and  $H_J^1$  are Toeplitz matrices that have zeros everywhere except for the first upper and first lower diagonals, which are composed of all ones. The subscript denotes the order of the matrices.

The structure of the potential matrix  $\mathcal{A}$  includes both past (left and up) and future (right and down) pixels for prediction of a pixel value. Such a representation precludes recursive computations as KBF or RTS filtering. We derive one-sided-forward representations by upper Cholesky factorization of  $\mathcal{A} = \mathcal{L}^T \mathcal{L}$  where the matrix  $\mathcal{L}$  is a lower triangular  $(IJ \times IJ)$  block bidiagonal matrix

$$\mathcal{L} = I_I \otimes L_i + K_I^1 \otimes Q_i \quad (4)$$

with  $(J \times J)$  lower and upper triangular blocks  $(L_i, Q_i)$ ,  $(1 \leq i \leq I)$ . The block  $K_I^1$  is an  $(I \times I)$  matrix that has zeros everywhere except for the first lower diagonal. Substituting in  $\mathcal{A}\vec{X} = \vec{e}$ , left multiplying by  $\mathcal{L}^{-T}$ , and expanding in terms of vectors  $X_i$  representing pixel values of row  $i$  gives

$$X_1 = L_1^{-1} \vec{w}_1 \quad X_i = -(L_i^{-1} Q_{i-1}) X_{i-1} + L_i^{-1} \vec{w}_i, \quad \text{for } 2 \leq i \leq I \quad (5)$$

where  $\vec{w} = L^{-T} \vec{e}$  represents the whitened field  $w(i, j)$  arranged in a row-major order. Using the relationship  $\vec{w} = L^{-T} \vec{e}$ , it follows that the noise field  $w(i, j)$  is white given that the covariance matrix of  $\vec{e}$  is  $\Sigma_e = \sigma^2 \mathcal{A}^{-1}$ . The forward regressors  $L_i$  and  $Q_i$  in (5) are evaluated by solving a Riccati type equation that follows by equating the main block diagonals and upper block diagonals in  $\mathcal{A} = \mathcal{L}^T \mathcal{L}$ . The resulting expressions are

$$S_I = B \quad S_i = (B - C S_{i+1}^{-1} C), \quad \text{for } (I-1) \geq i \geq 1 \quad (6)$$

where  $L_i^T L_i = S_i$  and  $Q_i = L_{i+1}^{-T} C$ . With real data fields and in actual applications, evaluation of the regressors  $L_i$  and  $Q_i$  is not required for all rows. In fact, the regressors converge asymptotically to a steady-state solution at a geometric rate [4];

<sup>1</sup>For convenience of notation, we express our results in terms of zero Dirichlet BCs. The results are generalizable to other BCs. In fact, we use the symmetric Neumann BC in our experiments.

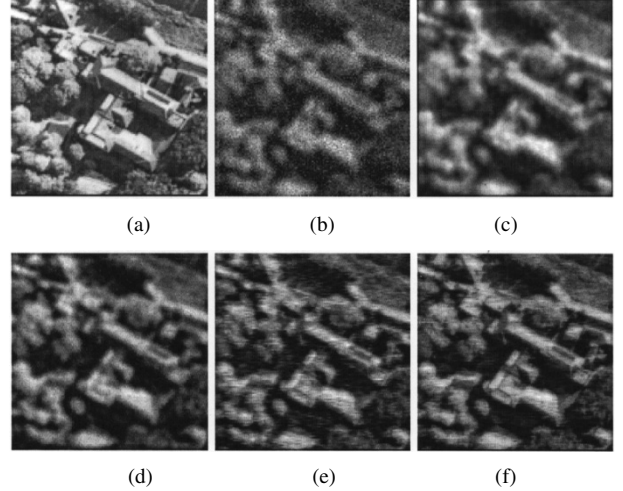


Fig. 2. Image restoration of the aerial image. (a) Original. (b) Noisy and blurred with truncated Gaussian blur (MSE = 1181.60). (c) Restored image with  $(3 \times 3)$  spatial averaging (MSE = 1015.3). (d) Restored image with Wiener filter (MSE = 439.6). (e) Restored image with RTS algorithm using causal prediction (MSE = 407.89). (f) Restored image with RTS algorithm using noncausal prediction (MSE = 207.54).

hence, only a few of the regressors need to be computed. For a first-order GMRP, the steady-state solution for the regressors is

$$S_\infty = \frac{B}{2} + \sqrt{\left(\frac{B}{2}\right)^2 - \beta_v^2 I_I}$$

with

$$L_\infty^T L_\infty = S_\infty \quad Q_\infty = L_\infty^{-T} C. \quad (7)$$

The notation  $\sqrt{G}$  defines the principal square root of a square matrix  $G$  such that  $\sqrt{G} \cdot \sqrt{G} = G$ . To save on the computational complexity, the regressors  $L_i$  and  $Q_i$  are approximated by their steady-state values,  $L_\infty$  and  $Q_\infty$ , in state (6). To use (6), we need estimates of the vertical and horizontal interactions,  $\beta_v$  and  $\beta_h$ . The evaluation of  $\beta_v$  and  $\beta_h$  is based on maximizing the likelihood function (see [4]), which is computationally intensive. To simplify the restoration procedure, we use the following approximations

$$\beta_v = \frac{\xi \chi_v}{|\chi_h| + |\chi_v|} \quad \beta_h = \frac{\xi \chi_h}{|\chi_h| + |\chi_v|} \quad (8)$$

where  $\chi_v$  and  $\chi_h$  denote the vertical and horizontal sample correlations of the GMRP and are given by

$$\chi_v = \frac{1}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=1}^N x(i, j) x(i+1, j)$$

$$\chi_h = \frac{1}{N(N-1)} \sum_{i=1}^N \sum_{j=1}^{N-1} x(i, j) x(i, j+1). \quad (9)$$

The term  $\xi$  is a positive tolerance and is bounded [4] by  $\xi < 1/(2 \cos \pi/N + 1)$ .

### III. BLURRING MODEL

We consider restoring images blurred by either the truncated Gaussian blur point spread function (psf)

$$g(i, j) = \begin{cases} k_1 e^{-(i^2+j^2)/2\sigma^2}, & |i-j| \leq 2 \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

or the out-of-focus blur psf

$$g(i, j) = \begin{cases} k_2, & i^2 + j^2 \leq 5 \\ 0, & \text{otherwise} \end{cases} \quad (11)$$

and corrupted by white Gaussian noise. The values of the constants,  $k_1$  and  $k_2$  in the psfs are selected such that  $\sum_{i,j} g(i, j) = 1$  in each blurring function. The constant  $\sigma$  in the truncated Gaussian blur corresponds to the standard deviation of the blur. In our experiments,  $\sigma$  is set equal to 6.

For the two psfs, in the matrix-vector notation, (1) is

$$Z_i = (G_1 X_{i-2} + G_2 X_{i-1} + G_3 X_i + G_2 X_{i+1} + G_1 X_{i+2}) + V_i, \quad \text{for } 1 \leq i \leq I \quad (12)$$

where two vectors  $X_i$  and  $Z_i$  represent row  $i$  of the original  $x(i, j)$  and blurred image  $z(i, j)$ . The observation noise vector  $V_i$  is obtained by lexicographic ordering of row  $i$  of the noise field  $v(i, j)$ . The blocks  $G_i$ 's are sparse and Toeplitz with exactly five nonzero diagonals for blur models (10) and (11).

#### A. Dynamic Models

The state model and the observation model (12) cannot be used directly in the RTS algorithm because of the difference in the dimensions of the field vectors  $X_i$ 's containing the state. Defining a new state vector  $\Psi_i = [X_{i-2}^T \ X_{i-1}^T \ X_i^T \ X_{i+1}^T \ X_{i+2}^T]^T$ , results in the following state and observation models:

$$\Psi_{i+1} = \underbrace{\begin{bmatrix} 0 & I_J & 0 & 0 & 0 \\ 0 & 0 & I_J & 0 & 0 \\ 0 & 0 & 0 & I_J & 0 \\ 0 & 0 & 0 & 0 & I_J \\ 0 & 0 & 0 & 0 & -L_\infty^{-1} Q_\infty \end{bmatrix}}_A \Psi_i + \underbrace{\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ L_\infty^{-1} \end{bmatrix}}_B W_i \quad (13)$$

$$Z_{i+1} = \underbrace{[G_1 \ G_2 \ G_3 \ G_2 \ G_1]}_G \Psi_{i+1} + V_i \quad (14)$$

which are used in the image restoration algorithm explained next. The state and observation noise are White and Gaussian, given by  $W_i \sim \mathcal{N}(0, Q = \sigma_w^2 I)$  and  $V_i \sim \mathcal{N}(0, R = \sigma_v^2 I)$ .

#### IV. IMAGE RESTORATION

The procedure for obtaining a restored image has three sequential stages that are outlined below.

- 1) *Parameter Estimation*: After subtracting the global mean, the horizontal ( $\beta_h$ ) and vertical ( $\beta_v$ ) field interactions are estimated. Based on the values of the interactions ( $\beta_h, \beta_v$ ), the asymptotic approximation of the regressors ( $L_i, Q_i$ ) is computed using (7).
- 2) *Steady-State Error Covariance Approximation*: Since the state and observation equations (13) and (14) are shift-invariant, the predictor covariance matrix  $P_{i+1|i}$  is approximated with its steady-state value (say  $P^{(p)}$ ) computed using the Riccati equation in the KBF

$$P_{i+1|i} = A P_{i|i} A^T + B Q B^T \quad (15)$$

$$P_{i+1|i+1} = [I - K_{i+1} G] P_{i+1|i} \quad (16)$$



Fig. 3. Image restoration of the Lena image. (a) Original. (b) Noisy and Blurred with out-of-focus blur (MSE = 501.33). (c) Restored image with  $(3 \times 3)$  spatial averaging (MSE = 432.24). (d) Restored image with Wiener filter (MSE = 305.13). (e) Restored image with RTS algorithm using causal prediction (MSE = 254.54). (f) Restored image with RTS algorithm using noncausal prediction (MSE = 166.46).

where

$$K_{i+1} = P_{i+1|i} G^T [G P_{i+1|i} G^T + R]^{-1}. \quad (17)$$

The initial condition is  $P_{1|0} = \sigma^2 I$ . The Frobenius norm of the difference  $(P_{i+2|i+1} - P_{i+1|i})$  is used as the convergence criterion for the predictor covariance matrix.

- 3) *RTS Smoothing*: The state model (13) and the observation model (14) with the state matrices provided by step 1) and the predictor covariance matrix given by step 2) are used as the basis for the double-sweep RTS smoother [1]. The forward sweep recursively computes the predictor estimate [(18)] ( $\hat{\Psi}_{i+1|i}$ ) and the filter estimate [(19)] ( $\hat{\Psi}_{i+1|i+1}$ ) using the KBF

$$\hat{\Psi}_{i+1|i} = A \hat{\Psi}_{i|i} \quad \text{with } \hat{\Psi}_{1|0} = 0 \quad (18)$$

$$\hat{\Psi}_{i+1|i+1} = \hat{\Psi}_{i+1|i} + K(Z_{i+1} - G \hat{\Psi}_{i+1|i}) \quad (19)$$

for  $1 \leq i \leq (I - 1)$ . Both Kalman gain  $K_i$  and the filter covariance matrix  $P_{i+1|i+1}$  converge and do not need any further updating during the KBF iterations. The steady-state values of the Kalman gain and filter covariance matrix are denoted by  $K$  and  $P^{(f)}$ . The predicted field  $\hat{\Psi}_{i+1|i}$  and the filtered field  $\hat{\Psi}_{i+1|i+1}$  are the outputs of this stage.

The backward sweep computes the smoother estimate [(20)]  $\hat{\Psi}_i$  from the predictor estimate ( $\hat{\Psi}_{i+1|i}$ ) and the filter estimate ( $\hat{\Psi}_{i+1|i+1}$ ) provided by the KBF

$$\hat{\Psi}_i = \hat{\Psi}_{i|i} + S(\hat{\Psi}_{i+1} - \hat{\Psi}_{i+1|i}) \quad \text{with } \hat{\Psi}_I = \hat{\Psi}_{I|I} \quad (20)$$

for  $(I - 1) \geq i \geq 1$ . The smoother gain  $S$  uses the steady-state values for the error covariances and is given by  $S = P^{(f)} A^T P^{(p)-1}$  where  $P^{(f)}$  is the steady-state value of the filter covariance matrix. The smoother covariance matrix  $P_i$  is obtained from the following iteration:

$$P_i = P^{(f)} + A(P_{i+1} - P^{(p)})A^T \quad \text{with } P_I = P^{(f)}. \quad (21)$$

TABLE I  
QUANTITATIVE COMPARISONS BASED ON PSNR AND MSE FOR  
RESTORATION SCHEMES USED IN THE LETTER

Restoration Scheme	Aerial Image		Lena Image	
	MSE	PSNR (dB)	MSE	PSNR (dB)
Noisy/Blurred Image	1181.6	17.41	501.33	21.13
Spatial Averaging	1015.3	18.07	432.24	21.77
Wiener Filtering	439.6	21.7	305.13	23.28
RTS w/ Causal Prediction	407.9	22.03	254.54	24.07
RTS w/ Noncausal Prediction	207.6	24.96	166.46	25.92

The smoothed image  $\hat{x}(i, j)$  obtained from the state vector  $\hat{\Psi}_i$  is the output of the final stage.

## V. NUMERICAL EXPERIMENTS

In this section, we compare the performance of the RTS-based smoothing algorithm that uses noncausal predictive models versus other enhancement schemes. We carry out qualitative and quantitative studies. The qualitative analysis is based on the subjective evaluation of the image restored by the selected schemes. The quantitative comparison is based on the means-square error (MSE) given by

$$\text{MSE} = \frac{1}{IJ} \sum_{i=1}^I \sum_{j=1}^J (x(i, j) - \hat{x}(i, j))^2 \quad (22)$$

where  $\hat{x}(i, j)$  denotes the restored version of the original image  $x(i, j)$ . We compare our RTS-based noncausal algorithm with the following restoration techniques.

- i) *Wiener Filter*: Using the orthogonality principle, the transfer function  $W(f_1, f_2)$  of the Wiener filter is

$$W(f_1, f_2) = \frac{G^*(f_1, f_2)S_{xx}(f_1, f_2)}{|G(f_1, f_2)|^2 S_{xx}(f_1, f_2) + S_{vv}(f_1, f_2)} \quad (23)$$

where  $S_{xx}(f_1, f_2)$  and  $S_{vv}(f_1, f_2)$  are the power spectral densities (PSDs) of the original image and additive noise. The term  $G(f_1, f_2)$  is the 2-D Fourier transform of the psf  $g(i, j)$  as defined in (1).

- ii) *Spatial Averaging*: The input pixel is replaced by a spatial average of its neighborhood pixels. A  $(3 \times 3)$  window is used as the neighborhood for the spatial averaging filter in our experiments.
- iii) *RTS with causal prediction*: This resembles the RTS algorithm described in the letter except for the prediction model, which is one-sided (causal) and is given by

$$\hat{x}(i, j) = \beta_d^c x(i-1, j-1) + \beta_v^c x(i-1, j) + \beta_h^c x(i, j-1) \quad (24)$$

where  $\beta_d^c$ ,  $\beta_v^c$ , and  $\beta_h^c$  are the diagonal, vertical, and horizontal field interactions of a third-order Markov mesh. The superscript denotes causal prediction.

- iv) *RTS with noncausal GMRP prediction* is the algorithm described in the letter.

Fig. 2 illustrates the experimental results of the restoration schemes [i)–iv)] with Neumann boundary conditions for an aerial test image distorted by the truncated Gaussian blur [(10)]

with  $\sigma$  set to 6. Gaussian noise with an SNR of 10 dB is added to the blurred image. The resulting distorted image is shown in Fig. 2(b) where for reference we also include the original image in Fig. 2(a). The outputs of the spatial averaging filter is shown in Fig. 2(c). The restored image does not include any additional postprocessing of images after restoration, as is the case for the later results. Since the spatial filter do not take the blurring model in consideration, it is unable to remove the distortions introduced by blur. Fig. 2(d)–(f) illustrates the outputs from the Wiener and RTS filter. In Fig. 2(d), we use the Wiener filter. Fig. 2(e) is obtained from the causal model of (24), while in Fig. 2(f), we use the noncausal GMRP model, described in Section II. The RTS filter based on the noncausal GMRP model [scheme iv)] exhibits the best performance, restoring important features such as edges distinctly than its counterpart based on the third-order causal Markov mesh [scheme iii)], which includes undesired horizontal streaking. The Wiener filter is computationally intensive, as it requires calculating the 2-D Fourier transform (and its inverse) of the blurred image, yet it does not clearly restore most of the features present in the aerial image. A second comparison based on the Lena image is shown in Fig. 3. The original image is distorted with the out-of-focus blur. Restoration results reinforce our earlier conclusions. The MSEs included in Table I highlight the quantitative improvement achieved by the noncausal RTS over schemes [i)–iii)] for the two test images.

## VI. SUMMARY

The letter presents a practical implementation of the RTS filter based on noncausal GMRP prediction model for restoration of blurred image corrupted by additive noise. We exploit the shift-invariant characteristics of the state matrices in the noncausal GMRP predictive model and use the steady-state solution of the Riccati equation in the RTS filter. The resulting implementation is computationally practical. Our experimental results outperform the Wiener filter and illustrate the superiority of the noncausal GMRP prediction model used in the RTS filter over a causal prediction model.

## REFERENCES

- [1] H. E. Rauch, F. Tung, and C. T. Striebel, "Maximum likelihood estimates of linear dynamic systems," *J. Amer. Inst. Aeronaut. Astronauts*, vol. 3, no. 8, pp. 1445–50, Aug. 1965.
- [2] S. Critin and M. R. Azimi-Sadjadi, "A full-plane block Kalman filter for image restoration," *IEEE Trans. Image Processing*, vol. 1, pp. 488–495, Oct. 1992.
- [3] R. Molina, A. K. Katsaggelos, J. Mateos, A. Heros, and C. A. Segall, "Restoration of severely blurred high range images using stochastic and deterministic relaxation algorithms in compound Gauss-Markov random fields," *Patt. Recognit.*, vol. 33, pp. 555–71, 2000.
- [4] J. M. F. Moura and N. Balram, "Recursive structure of noncausal Gauss Markov random fields," *IEEE Trans. Inform. Theory*, vol. 38, pp. 334–354, Mar. 1992.
- [5] A. Asif and J. M. F. Moura, "Inversion of block matrices with  $L$ -block banded inverse," Dept. Comput. Sci., York Univ., Toronto, ON, Canada, 2002, submitted for publication.
- [6] N. Balram and J. M. F. Moura, "Recursive enhancement of noncausal images," in *Proc. IEEE Int. Conf. Acoustics, Speech Signal Processing*, May 1991, pp. 2997–3000.
- [7] A. Tekalp, H. Kaufman, and J. Woods, "Identification of image and blur parameters for the restoration of noncausal blurs," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, vol. 10, Apr. 1985, pp. 656–59.