Mathematical Analysis of Super-Resolution Methodology

Michael K. Ng and Nirmal K. Bose

An image acquisition system composed of an array of sensors, where each sensor has a subarray of sensing elements of suitable size, has recently been popular for increasing the spatial resolution with high signal-to-noise ratio beyond the performance bound of technologies that constrain the manufacture of imaging devices. The attainment of super resolution (SR) from a sequence of degraded undersampled images could be viewed as reconstruction of the high-resolution (HR) image from a finite set of its projections on a sampling lattice. This can then be formulated as an optimization problem whose solution is obtained by minimizing a cost function. The approaches adopted and their analysis to solve the formulated optimization problem are crucial, and the subsequent documentation will trace the issues leading to cutting-edge research in the subject.

The image acquisition scheme is important in the modeling of the degradation process. The need for model accuracy is undeniable in the attainment of SR along with the design of the algorithm whose robust implementation will produce the desired quality in the presence of model parameter uncertainty. To keep the presentation focused and of reasonable size, data acquisition with multisensors instead of, say, a video camera is considered. Multiple undersampled images of a scene are often obtained by using multiple identical image sensors which are shifted relative to each other by subpixel displacements [8], [21], [22]. The resulting HR image reconstruction problem using a set of low-resolution (LR) images captured by the image sensors is interesting because it is closely related to the design of high-definition television (HDTV) and very high-definition (VHD) image sensors. Charge-coupled device (CCD) image sensor arrays, where each sensor consists of a rectangular subarray of sensing elements, produce discrete images whose sampling rate and resolution are determined by the physical size of the sensing elements. If multiple CCD image sensor arrays are shifted relative to each other by exact subpixel values, the reconstruction of HR images is sometimes modeled [8] by

$$\bar{\mathbf{g}} = H\mathbf{f} \quad \text{and} \quad \mathbf{g} = \bar{\mathbf{g}} + \eta, \tag{1}$$

where $\mathbf{f}$ is the desired HR image, H is the blur matrix, and $\bar{\mathbf{g}}$ is the observed HR image by interlacing the LR images from sensors. η is the additive, possibly Gaussian, noise. Figure 1 shows the method of forming a 4×4 image $\mathbf{g}$ with a 2×2 sensor array where each $\mathbf{g}_{ij}$ has a 2×2 sensing elements, i.e., four 2×2 LR images. Here the blur matrix is constructed from the averaging of the pixel values. As an example, Figure 2(a) illustrates that the point spread func-

1053-5888/03/$17.00©2003IEEE

tion of the sensor applies to the image by averaging of the four pixel values. Since the system described in (1) is ill-conditioned, a solution for $\mathbf{f}$ is constructed by applying a regularization technique that involves a functional $\mathcal{R}(\mathbf{f})$, which captures the regularity in $\mathbf{f}$, and a tuning (regularization) parameter α that controls the degree of regularity of the solution to the minimization problem:

$$\min_{\mathbf{f}}\left\{\|H\mathbf{f}-\mathbf{g}\|_2^2+\alpha\mathcal{R}(\mathbf{f})\right\}. \tag{2}$$

The maximum a posteriori (MAP) regularization technique [8], [39], the L_2-norm regularization functional $\mathcal{R}(\mathbf{f})=\|\mathbf{f}\|_2^2$, and the H_1-norm regularization functional $\mathcal{R}(\mathbf{f})=\|L\mathbf{f}\|_2^2$ [36], where L is the discretization matrix of the first order differential operator, have been considered and used in HR image reconstruction.

The main task is to recover the HR image $\mathbf{f}$ given the blurring matrix H and a blurred and noisy image $\mathbf{g}$ using regularization models. There are several problems that must be handled and solved to efficiently obtain a high quality HR image.

▲ Owing to the blurring (convolution) process, the boundary values of g are not completely determined by the original HR image f inside the scene. They are also affected by the values of f outside the scene. How can we handle boundary conditions?

▲ Iterative methods can be applied to solve the minimization problem (2). However, due to the ill-conditioning of the blurring matrix, the convergence of iterative methods can be very slow. How can we speed up the convergence of iterative methods?

▲ Usually, we assume that subpixel displacements are known. However, this assumption may not hold in practice. How can we estimate subpixel displacements?

▲ In some situations, we do not have enough LR images to resolve the HR image. How can we modify the regularization model to deal with this situation?

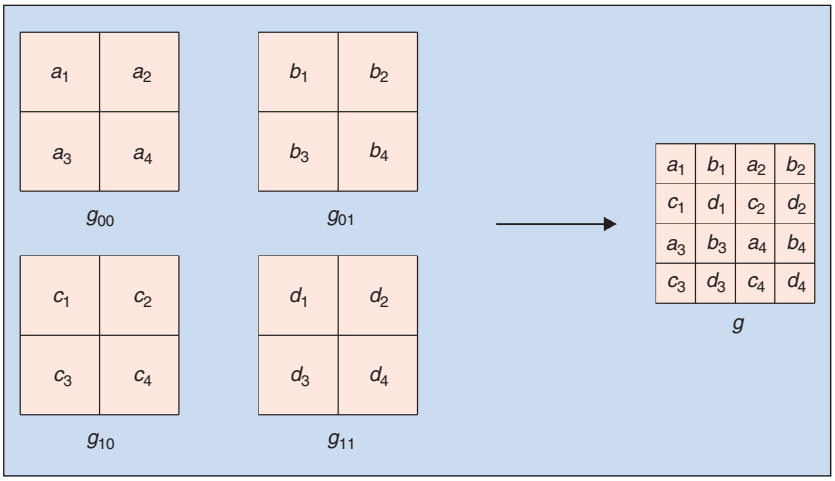
▲ The H_1-norm regularization tends to attenuate the high frequency information of the HR image. If we consider edge-preserving regularization models, can we design an efficient iterative method for solving the corresponding minimization problem?

In the following sections, we will study the above-mentioned problems and provide effective and efficient methodologies to solve these problems.

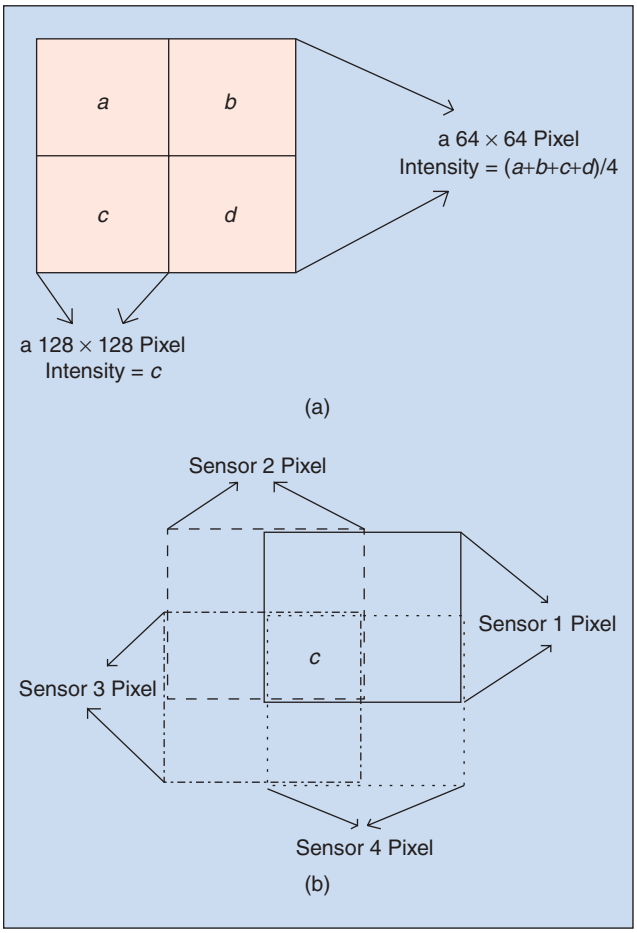
Image Boundary

The boundary values of the observed image $\mathbf{g}$, because of the blurring process, are affected by the values of original image $\mathbf{f}$ outside the scene. Figure 2(b) illustrates that the application of the point spread function of the sensor to the image involves the pixel values outside the scene. When solving for $\mathbf{f}$ from (2), some assumptions on the values of $\mathbf{f}$ outside the scene are needed. These assumptions are linked to boundary conditions [1], [16].

The traditional choice was the imposition of zero boundary condition outside the scene [8], i.e., a dark background was assumed outside the scene. The blur matrix, in this case, is a block-Toeplitz with Toeplitz-block (BTTB) matrix [35], which occurs naturally in many



▲ *1. Formation of the observed HR image.*



▲ *2. Example of blur operator.*

applications. When the blur is of the linear shift-invariant type, the image restoration problem is often modeled as a system of linear equations, where the characterizing matrix has a structure that can be very ill-conditioned. An effective way to combat this problem is to approximate the BTTB matrix by means of a block-circulant with circulant-block (BCCB) matrix. The sequences of BTTB and BCCB matrices of increasing size are asymptotically equivalent in a certain sense and the BCCB matrix is a very natural and well-used preconditioner [6].

For practical applications, one often resorts to periodic boundary conditions [1], [16]. The blur matrix, in this case, is a BCCB matrix which can be diagonalized and inverted by the efficiently implementable two-dimensional (2-D) discrete Fourier transform (DFT) matrix. When the assumptions of either zero or periodic boundary conditions are not satisfied by the captured images, however, ringing effects may occur at the boundary of the reconstructed image. Figure 3 shows the boundary artifacts of the reconstructed images by using the zero and periodic boundary conditions. The problem tends to be more severe if the image is reconstructed from data acquired by a large sensor array since the number of pixel values of the image affected by the sensor array increases.

The Neumann boundary condition on the image (the scene immediately outside is a reflection of the original scene at the boundary) is then assumed [3], [24], [28]. In this case, the blur matrix is a block-Toeplitz-plus-Hankel with Toeplitz-plus-Hankel-block (BTHTHB) matrix [35]. A symmetric BTHTHB matrix can be diagonalized by the DCT matrix, and because of this a BTTB matrix encountered in image reconstruction problems is often approximated by a symmetric BTHTHB matrix for computational efficiency in implementation. Very recently, the error caused by the approximation is analyzed and a simple modification of the observed image that can reduce the approximation error is proposed in [23]. Experimental results in [10] and [35] have shown that the Neumann image model gives better reconstructed HR images than those using the zero or periodic boundary conditions. In Figure 3, we see that the boundary artifacts of the reconstructed image using the Neumann boundary condition is very small. In [39], it has been shown that the use of the Neumann boundary condition leads to smaller expected errors in the reconstructed image when the original image is statistically stationary near the boundary.

Here we visualize the structures of blurring matrices based on these three different boundary conditions.

Zero boundary condition (BTTB matrix):

$$\begin{pmatrix} h_0 & \cdots & h_{-m} & & 0 \\ \vdots & \ddots & \ddots & \ddots & \\ h_m & \ddots & \ddots & \ddots & h_{-m} \\ & \ddots & \ddots & \ddots & \vdots \\ 0 & & h_m & \cdots & h_0 \end{pmatrix}$$

Periodic boundary condition (BCCB matrix):

$$\begin{pmatrix} h_0 & \cdots & h_{-m} & & 0 \\ \vdots & \ddots & \ddots & \ddots & \\ h_m & \ddots & \ddots & \ddots & h_{-m} \\ & \ddots & \ddots & \ddots & \vdots \\ 0 & & h_m & \cdots & h_0 \end{pmatrix} + \begin{pmatrix} 0 & & h_m & \cdots & h_1 \\ & & & \ddots & \vdots \\ h_{-m} & & & & h_m \\ \vdots & \ddots & & & \\ h_{-1} & \cdots & h_{-m} & & 0 \end{pmatrix}$$

Neumann boundary condition (BTHBTH matrix):

$$\begin{pmatrix} h_0 & \cdots & h_{-m} & & 0 \\ \vdots & \ddots & \ddots & \ddots & \\ h_m & \ddots & \ddots & \ddots & h_{-m} \\ & \ddots & \ddots & \ddots & \vdots \\ 0 & & h_m & \cdots & h_0 \end{pmatrix} + \begin{pmatrix} h_1 & \cdots & h_m & & 0 \\ \vdots & \ddots & & & \\ h_m & & & & h_{-m} \\ & & & \ddots & \vdots \\ 0 & & h_{-m} & \cdots & h_{-m} \end{pmatrix}.$$

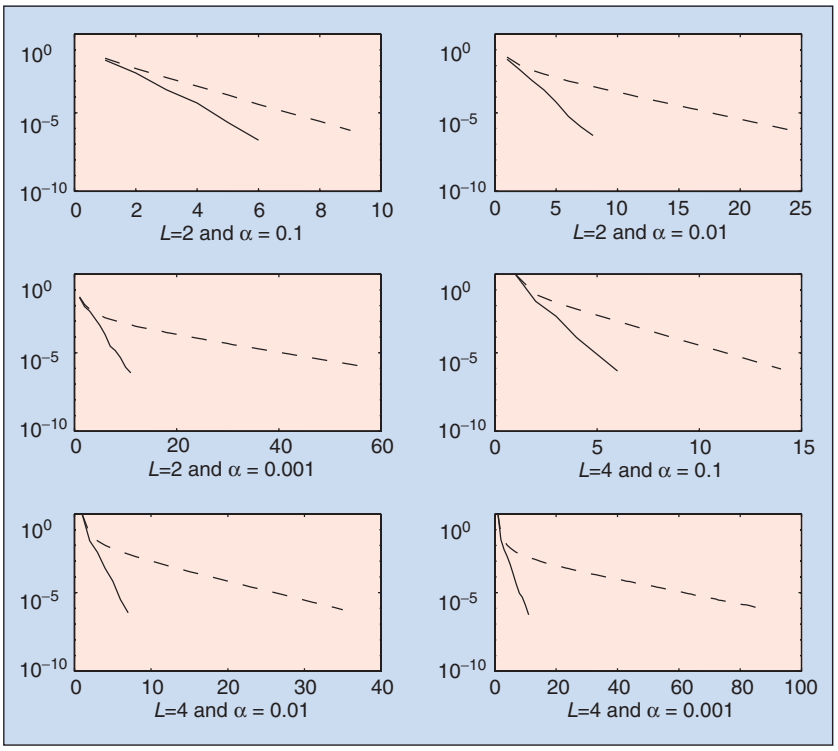
▲ *3. Reconstructed images using (a) the zero boundary condition, (b) the periodic boundary condition, and (c) the Neumann boundary condition.*

Fast Iterative Methods

In the case of Tikhonov regularization in (2), the linear system of equations to be solved is

$$\left(H^T H + \alpha R\right)\mathbf{f} = H^T \mathbf{g} \tag{3}$$

where R is the matrix associated with the regularization functional $\mathcal{R}(\mathbf{f})$ in (2). When the MAP regularization technique is used, $R = (I-S)^T (I-S)$, where S comes from a symmetric noncausal statistical model of the image [8], [39], [5]. With the Neumann boundary condition, this matrix S, resulting from a symmetric blurring function, is a symmetric BTHTHB matrix and, therefore, can be diagonalized by 2-D DCT [35], [39]. When $\mathcal{R}(\mathbf{f}) = \|\mathbf{f}\|_2^2$ or $\mathcal{R}(\mathbf{f}) = \|L\mathbf{f}\|_2^2$, the matrix R is either equal to the identity matrix or the discrete Laplacian matrix. Note that in all cases the matrix R can be diagonalized by the 2-D DCT matrix. Hence, inversion in (3) can be done using three two-dimensional (2-D) fast cosine transforms (one for finding the eigenvalues of the coefficient matrix and the remaining two for transforming the right-hand side and the solution vector). Thus the total cost of reconstruction of an n-by-n HR image requires $O(n^2 \log n)$ operations akin to what is required for the periodic boundary condition (characterized by BCCB matrix diagonalizable by 2-D DFT).



▲ *4. The convergence curves of using preconditioners and without using preconditioners for different sizes L of sensor array and regularization parameters α. The unbroken and broken lines are associated, respectively, with and without the use of preconditioners.*

Perfect subpixel displacements $\mathbf{e}$ are practically impossible to realize. Therefore, blur operators in multisensor HR image reconstruction are space variant. The blur matrix for the whole L-by-L sensor array is made up of blur matrices from each sensor:

$$H(\mathbf{e}) = \sum_{L-1}^{l_1=0}\sum_{L-1}^{l_2=0} D_{l_1 l_2} H_{l_1 l_2}(\mathbf{e}), \tag{4}$$

were $D_{l_1 l_2}$ is a diagonal matrix with diagonal element equal to 1 if the corresponding component of $\mathbf{g}$ comes from the (l_1, l_2)th sensor and zero otherwise. With the Tikhonov regularization, the linear system is

$$A\mathbf{f} \equiv \left(H^T(\mathbf{e})H(\mathbf{e}) + \alpha R\right)\mathbf{f} = H^T(\mathbf{e})\mathbf{g} \equiv \mathbf{b}. \tag{5}$$

The blurring matrix $H(\mathbf{e})$ has the same banded structure as that of H, but with some entries perturbed. It is a near BTHTHB matrix but it can no longer be diagonalized by the DCT. Therefore, the corresponding linear system may be solved by the preconditioned conjugate gradient method [32], [39].

Convergence Rate

The conjugate gradient (CG) method was invented in the 1950s [18] as a direct method for solving symmetric positive definite systems. It has come into wide use over the last 20 years as an iterative method. Let us consider (5) where the coefficient matrix A is a symmetric positive definite matrix and the right-hand side vector $\mathbf{b}$. Given an initial guess $\mathbf{f}_0$ and the corresponding initial residual $\mathbf{r}_0 = \mathbf{b} - A\mathbf{f}_0$, the kth iterate $\mathbf{f}_k$ of CG minimizes the functional $\Psi(\mathbf{f}) \equiv (1/2)\mathbf{f}^T A\mathbf{f} - \mathbf{f}^T\mathbf{b}$, over $\mathbf{f}_0 + \mathbb{K}_k$ where $\mathbb{K}_k$ is the kth Krylov subspace

$$\mathbb{K}_k \equiv \text{span}\left(\mathbf{r}_0, A\mathbf{r}_0, \ldots, A^{k-1}\mathbf{r}_0\right),\ k = 1,2,\ldots.$$

Note that if $\mathbf{f}^*$ minimizes $\Psi(\mathbf{f})$, then $\nabla\Psi(\mathbf{f}^*) = A\mathbf{f} - \mathbf{b} = 0$ and hence $\mathbf{f}^*$ is the solution.

Denote $\mathbf{f}^*$ the true solution of the system and define the norm

$$\|\mathbf{f}\|_A \equiv \sqrt{\mathbf{f}^T A\mathbf{f}}.$$

One can show that minimizing $\Psi(\mathbf{f})$ over $\mathbf{f}_0 + K_k$ is the same as minimizing $\|\mathbf{f} - \mathbf{f}^*\|_A$ over $\mathbf{f}_0 + K_k$. Since any $\mathbf{y} \in \mathbf{f}_0 + K_k$ can be written as

$$\mathbf{y} = \mathbf{f}_0 + \sum_{i=0}^{k-1} \beta_i A^i \mathbf{r}_0$$

for some coefficients $\{\beta_i\}_{i=0}^{k-1}$, we can express $\mathbf{f}^* - \mathbf{y}$ as

$$\mathbf{f}^* - \mathbf{y} = \mathbf{f}^* - \mathbf{f}_0 - \sum_{i=0}^{k-1} \beta_i A^i \mathbf{r}_0.$$

As $\mathbf{r}_0 = \mathbf{b} - A\mathbf{f}_0 = A(\mathbf{f}^* - \mathbf{f}_0)$, we have

$$\mathbf{f}^* - \mathbf{y} = \mathbf{f}^* - \mathbf{f}_0 - \sum_{i=0}^{k-1} \beta_i A^{i+1}(\mathbf{f}^* - \mathbf{f}_0) = p(A)(\mathbf{f}^* - \mathbf{f}_0),$$

where the polynomial

$$p(z) = 1 - \sum_{i=0}^{k-1} \beta_i z^{i+1}$$

has degree k and satisfies $p(0) = 1$. Hence

$$\left\| \mathbf{f}^* - \mathbf{f}_k \right\|_A = \min_{p \in \mathbb{P}_k,\, p(0)=1} \left\| p(A)(\mathbf{f}^* - \mathbf{f}_0) \right\|_A, \tag{6}$$

where $\mathbb{P}_k$ is the set of polynomials of degree k. Since symmetric positive definite matrices asserts that $A = U\Lambda U^*$, where U is a unitary matrix whose columns are the eigenvectors of A and Λ is the diagonal matrix with the positive eigenvalues of A on the diagonal. Since $UU^* = U^*U = I$, we have $A^k = U\Lambda^k U^*$. Hence $p(A) = Up(\Lambda)U^*$. Define $A^{\frac{1}{2}} = U\Lambda^{\frac{1}{2}}U^*$, we have

$$\left\| p(A)\mathbf{f} \right\|_A = \left\| A^{\frac{1}{2}} p(A)\mathbf{f} \right\|_2 \le \left\| p(A) \right\|_2 \left\| \mathbf{f} \right\|_A.$$

Together with (6), this implies that

$$\left\| \mathbf{f}^* - \mathbf{f}_k \right\|_A \le \left\| \mathbf{f}^* - \mathbf{f}_0 \right\|_A \min_{p \in P_k,\, p(0)=1} \max_{\lambda \in \sigma(A)} |p(\lambda)| \tag{7}$$

where $\sigma(A)$ is the set of eigenvalues of A.

The convergence rate of the conjugate gradient method has been well studied (see, for example, Golub and Van Loan [15]). It depends on the condition number $\kappa(A)$ of the matrix A and on how clustered the spectrum of A is. If the spectrum is not clustered, as is usually the case for blur matrices, a good estimate of the convergence rate is given in terms of the condition number

$$\kappa(A) = \frac{\text{the largest eigen value of } A}{\text{the smallest eigenvalue of } A}.$$

The convergence rate of the method can be expressed as

$$\frac{\left\| \mathbf{f}_k - \mathbf{f}^* \right\|_A}{\left\| \mathbf{f}_0 - \mathbf{f}^* \right\|_A} < 2\left(\frac{\sqrt{\kappa(A)} - 1}{\sqrt{\kappa(A)} + 1} \right)^k.$$

This indicates that the rate of convergence can be very slow if $\kappa(A)$ is large. Thus, the method will converge in a large number of iterations and hence the complexity of solving the system (5) is very large. Figure 4 shows the slow convergence of the conjugate gradient method when applied to solving the system (5) for different values of L and α.

One way to speed up the convergence rate of the method is to precondition the system. Thus, instead of solving (5), one solves the preconditioned system

$$P^{-1}\left(H^T(\mathbf{e})H(\mathbf{e}) + \alpha R \right)\mathbf{f} = P^{-1}H^T(\mathbf{e})\mathbf{g}. \tag{8}$$

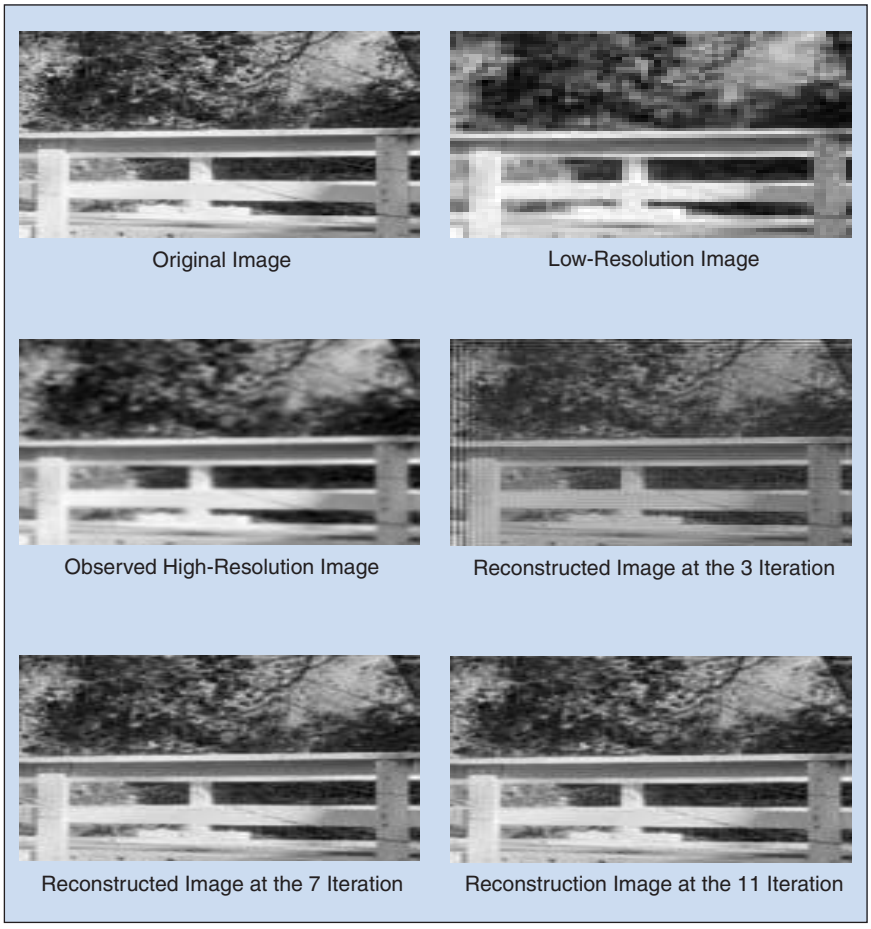
The symmetric positive definite matrix P, called the preconditioner, should be chosen according to the following criteria:

▲ P should be constructed very efficiently.

▲ $Px = y$ should be solved for x very efficiently.

▲ The spectrum of $P^{-1}(H^T(e)H(e) + \alpha R)$ should be clustered or its condition number should be reduced.

The first two criteria occur from the operation count per iteration as that is the count for the non-preconditioned system. The third criterion comes from the fact that the more



▲ *5. Reconstructed images using different PCG iterations.*

clustered the eigenvalues are (or the less the condition number is), the faster the convergence of the method will be (see, for instance, [27]).

Since $H^T H + \alpha R$ can be diagonalized by the DCT matrix, therefore $H^T H + \alpha R$ can be employed as the preconditioner for $H^T(\mathbf{e})H(\mathbf{e}) + \alpha R$ (the first two criteria). In [39], it was shown that when either the MAP regularization or the L_2-norm or the H_1-norm regularization functional is used, the eigenvalues of the preconditioned matrix $(H^T H + \alpha R)^{-1}(H^T(\mathbf{e})H(\mathbf{e}) + \alpha R)$ are clustered around the fixed point at 1 for sufficiently small subpixel displacement errors [32]. The conjugate gradient method, when applied to solving the preconditioned system (8), converges superlinearly [36]. Figure 4 shows the fast convergence of the conjugate gradient method when applied to solving the preconditioned system for different values of L and α. In Figure 5, we also show the reconstructed image can be obtained within a few PCG iterations. More precisely, for any given $\varepsilon > 0$, there exists a constant $c(\varepsilon) > 0$ such that the error vector $\mathbf{f}_k - \mathbf{f}^*$ of the preconditioned conjugate gradient method at the kth iteration satisfies

$$\frac{\|\mathbf{f}_k - \mathbf{f}^*\|_A}{\|\mathbf{f}_0 - \mathbf{f}^*\|_A} \le c(\varepsilon)\varepsilon^k. \tag{9}$$

In [32], it was further shown that the convergence rate of the conjugate gradient method for (8) depends linearly on the displacement errors arising from imperfect subpixel locations. Thus the cost per each iteration is $O(n^2 \log n)$ operations and hence the total cost for finding the HR image vector from (8) using the MAP, L_2-norm or H_1-norm regularization is still $O(n^2 \log n)$ operations.

We remark that the use of cosine transform-based matrices as preconditioners for HR image reconstruction problems allows the use of fast cosine transform throughout the computations. Notice that fast cosine transform is highly parallelizable and has been implemented on multiprocessors efficiently. Since the PCG method is easily parallelizable, too, the cosine transform-based preconditioned conjugate gradient method is well adapted for parallel computing [30].

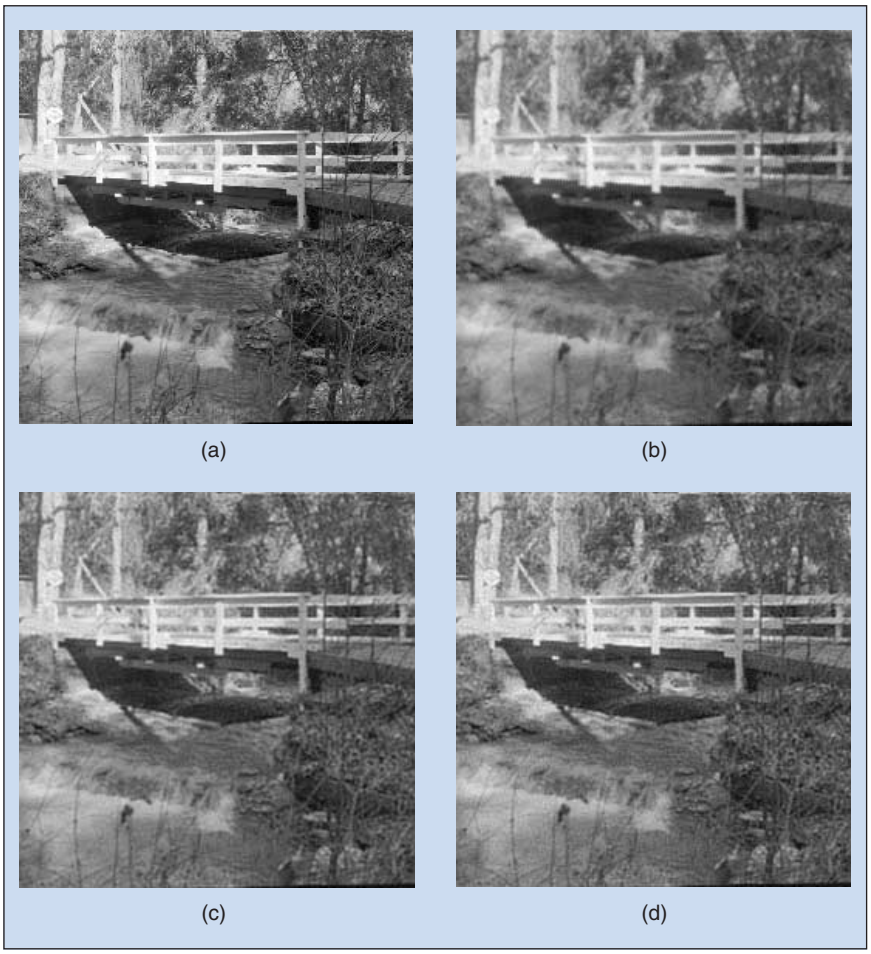
Regularized Constrained Total Least Squares

The spatial resolution of an image is often determined by imaging sensors. In a CCD camera, the image resolution is determined by the size of its photo-detector. An ensemble of several shifted images could be collected by a prefabricated planar array of CCD sensors and one may reconstruct with higher resolution. This is equivalent to an effective increase of the sampling rate by interpolation. Fabrication limitations are known to cause subpixel displacement errors, which, coupled with observation noise, limit the deployment of least squares techniques in this scenario. However, the displacement errors may not be known exactly. Total least squares (TLS) [43] is an effective technique for solving a set of such error-contaminated equations and, therefore, is an appropriate method for consideration in our HR image reconstruction applications. The TLS HR image reconstruction problem can be formulated as follows:

$$\min_{\mathbf{f},\ \text{arbitrary } V} \left\| [(H(\mathbf{e}) + V)\mathbf{f} - \mathbf{g}] \right\|_2^2 + \alpha \mathcal{R}(\mathbf{f}). \tag{10}$$

The solution of the minimization problem (10) can be determined by solving an eigenvalue problem [43].

A possible drawback of using such a conventional TLS approach is that the matrix V is arbitrary in the minimization process and the formulation in (10) is not constrained to handle special structure of the blur matrix



▲ *6. (a) Original image of size 256×256. (b) Observed blurred and noisy image 256×256; PSNR = 24.20 dB. (c) Reconstructed image by RLS; PSNR = 25.64 dB. (d) Reconstructed image by RCTLS; PSNR = 25.88 dB.*

The image acquisition scheme is important in the modeling of the degradation process.

from multisensors. We note that the spatial invariance of the blurring matrix for each sensor translates into the spatial invariance of the displacement error in the blur matrix. In [32], an image processing technique that leads to the deployment of constrained total least squares (CTLS) theory is described. The regularized constrained total least squares (RCTLS) solution to the problem given requires the minimization of a nonconvex and nonlinear cost functional:

$$J_{TLS}(\mathbf{f},\mathbf{e}) \equiv \min_{\mathbf{f},\mathbf{e}\ \text{such that}\ H(\mathbf{e})\ \text{is BTHBTH}} \|[H(\mathbf{e})\mathbf{f}-\mathbf{g}]\|_2^2 + \alpha\mathcal{R}(\mathbf{f}). \tag{11}$$

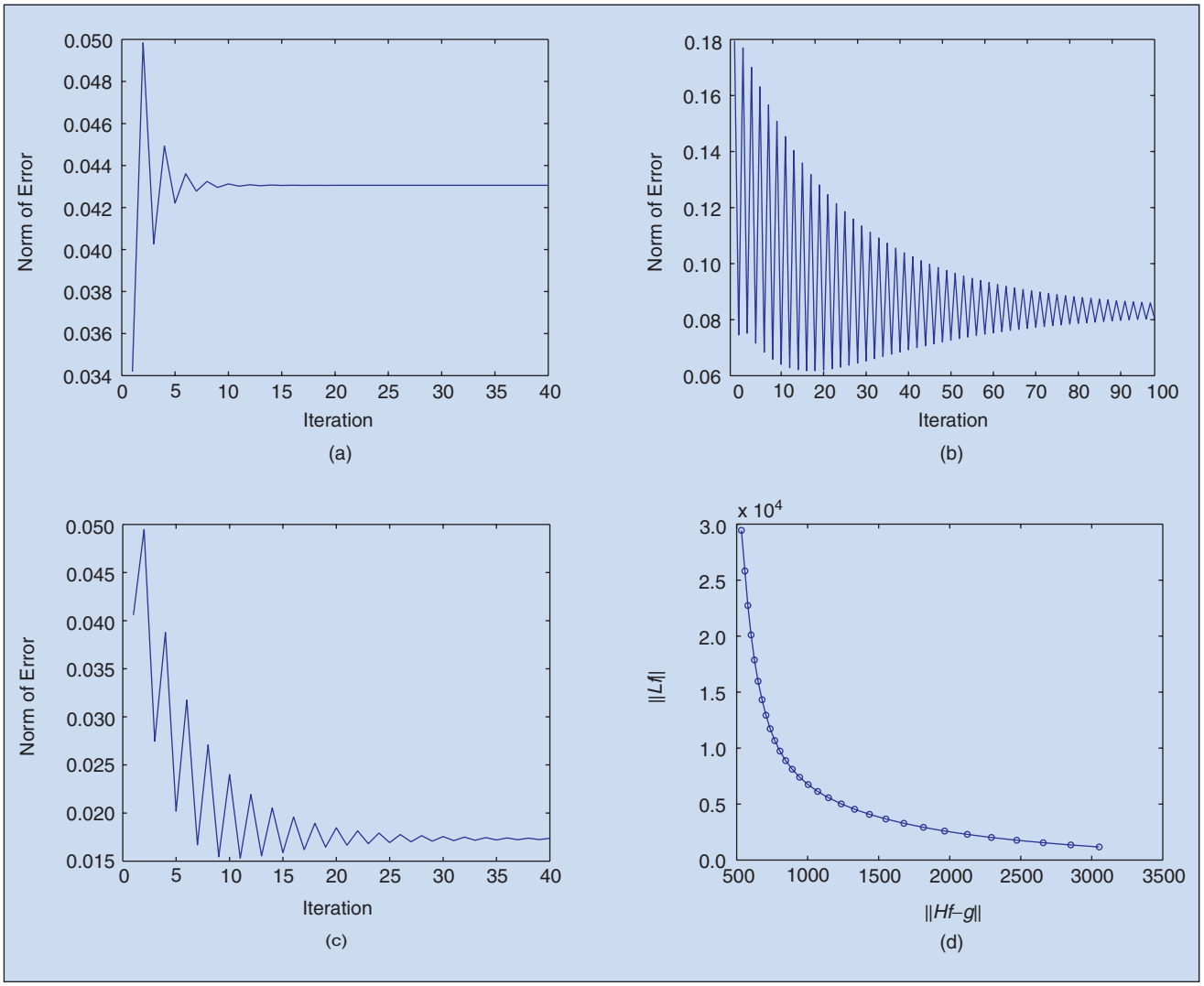
The solution of the minimization problem (11) cannot be determined by solving an eigenvalue problem. In [31], an iterative algorithm that takes advantage, computationally, of the fast solvers for image reconstruction problems with known displacement errors is developed. Before solving for $\mathbf{f}$ in the RCTLS formulation, it is first noted that for a given $\mathbf{f}$, the function $J_{\text{TLS}}(\mathbf{f},\cdot)$ is convex with respect to $\mathbf{e}$ (up to negligible terms), and for a given $\mathbf{e}$, the function $J_{\text{TLS}}(\cdot,\mathbf{e})$ is also convex with respect to $\mathbf{f}$. Therefore, with an initial guess $\mathbf{e}_0$, one can minimize (11) by first solving

$$J_{\text{TLS}}(\mathbf{f}_1,\mathbf{e}_0) = \min_{\mathbf{f}} J_{TLS}(\cdot,\mathbf{e}_0)$$

and then

$$J_{\text{TLS}}(\mathbf{f}_1,\mathbf{e}_1) = \min_{\mathbf{e}_1} J_{TLS}(\mathbf{f}_1,\cdot).$$

We note that the first subproblem can be solved by using the preconditioned conjugate gradient method, and the second subproblem requires solving a small least squares problem. Therefore, both subproblems can be solved very efficiently. An alternating minimization algorithm is developed in which the function value $J_{\text{TLS}}(\mathbf{f}_n,\mathbf{e}_n)$ always decreases as n increases. In most cases, the algorithm



▲ *7. The norm of the errors with (a) $\alpha = 1\times10^{-4}$, (b) $\alpha = 1$, (c) $\alpha = \alpha_{opt}$, and (d) the L-curve.*

converges to a local minimizer. It is interesting to note that the convergence analysis of this algorithm is still under investigation. The reconstructed image by regularized least squares (RLS) and the reconstructed image by RCTLS are shown in Figure 6(c) and (d), respectively. The reconstructed HR image using RCTLS shows improvement both in image quality and PSNR.

In the proposed RCTLS algorithm, we find that the choice of "proper" regularization parameter α is very important. In Figure 7, we plot the norm of the error between the real subpixel displacements and the estimated subpixel displacements by the algorithm for different values of α. The speed of convergence decreases as α increases. In the case when $\alpha = 10^{-4}$ [Figure 7(a)] and $\alpha = 1$ [Figure 7(b)], the norm of error after convergence is greater than that in the first iteration. The "inappropriate" value of α makes the RCTLS solution fall into a local minimum. The L-curve method [17] may be used to get the optimum value of the regularization parameter. The L-curve method to estimate the "proper" α for RCTLS is used here. The L-curve plot is shown in Figure 7(d). With α_{opt} retrieved by L-curve, we see in Figure 7(c) that the RCTLS converges to a better minimum point (the norm of error is significantly smaller than those obtained by choosing $\alpha = 10^{-4}$ and $\alpha = 1$).

Color Images

Multispectral restoration of a single image is a three-dimensional (3-D) reconstruction problem where the third axis incorporates different wavelengths. We are interested in color images because there are many applications. Color image can be regarded as a set of three images in their primary color channels (red, green, and blue). Monochrome processing algorithms applied to each channel independently are not optimal because they fail to incorporate the spectral correlation between the channels. Under the assumption that the spatial intrachannel and spectral interchannel correlation functions are product separable, Hunt and Kubler [20] showed that a multispectral (e.g., color) image can be decorrelated by the Karhunen-Loeve transform (KLT). After decorrelating multispectral images, the Wiener filter can be applied independently to each channel, and the inverse KLT gives the restored color image. It has recently been shown [33] that single color image restoration with multisensors can be formulated as a regularized least squares problem that can be solved efficiently using the fact that the DCT can diagonalize the linear system of equations (resulting from the use of Neumann boundary conditions), characterized by a BTHTHB matrix.

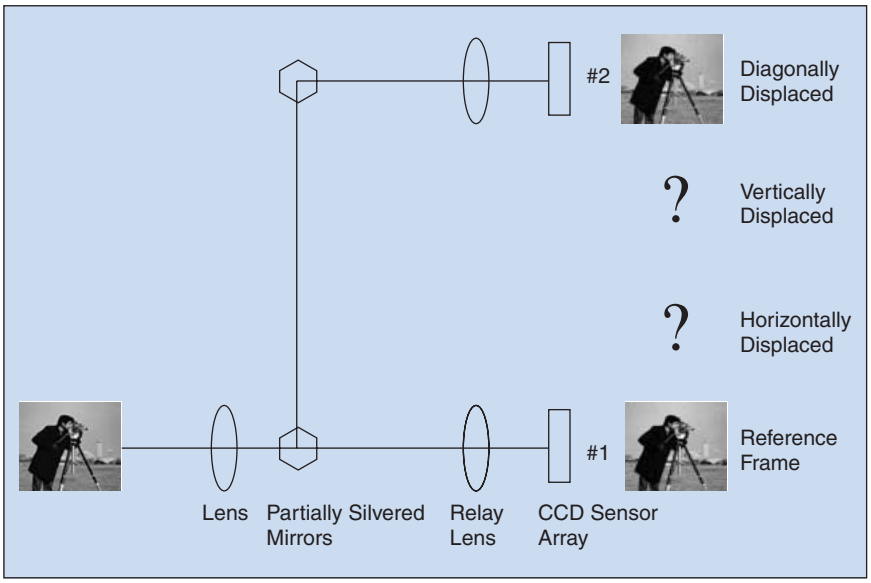
The tradeoff between the size of the scene and its resolution has to be addressed.

There is considerable scope for incorporating regularization methods like the L-curve [7] in this approach for further improvement in quality of restoration both in the case of a single image as well as multispectral video sequences. In [34], Ng et al. also extended the HR image reconstruction method to multiple undersampled color images. The key issue is to employ the cross-channel regularization matrix to capture the changes of reflectivity across the channels.

Insufficient LR Images

Previously, it was assumed that enough LR images are available to resolve the HR image. Here, the situation where some LR images are missing is considered. For example, in a 2×2 sensor array, we should have the reference, the horizontally displaced, the vertically displaced, and the diagonally displaced LR image to resolve the HR image. In Figure 8, we consider the situation that we only have the reference and the diagonally displaced LR image. The other two displaced LR images are missing.

We note that the optimization problem in (2) can be used to reconstruct HR images. However, since the blur matrix is no longer BTHTHB, therefore DCT-based preconditioners do not work well [38]. Figure 9 shows the slow convergence of the preconditioned conjugate gradient method. We find that the preconditioner is not effective. In [37], Ng et al. considered the following joint minimization problem:



▲ *8. Missing LR images.*

Due to the ill-conditioning of the blurring matrix, the convergence of iterative methods can be very slow. How can we speed up the convergence?

$$\min_{\mathbf{f},u} J_{IL}(\mathbf{f},\mathbf{u}) \equiv \min_{\mathbf{f},\mathbf{u}} \frac{1}{2}\|H(\mathbf{e})\mathbf{f} - (\mathbf{u}+\mathbf{v})\|_2^2 + \alpha_1 \mathcal{R}(\mathbf{f}) + \alpha_2 \widetilde{\mathcal{R}}(\mathbf{u}) \quad (12)$$

where $\mathbf{u}$ corresponds to the missing observed image pixels because of the missing or faulty sensors in the sensor array and $\mathbf{u}+\mathbf{v}$ and H are the full observed HR image and degradation matrix, respectively, by combining all the sensors in the sensor array. Here $\mathcal{R}$ is the regularization functional that was discussed earlier, and $\widetilde{\mathcal{R}}$ is the regularization functional for $\mathbf{u}$, α_1 and α_2 are positive parameters which measure the tradeoff between a good fit and the regularity of the solutions $\mathbf{f}$ and $\mathbf{u}$. Due to the local averaging of the pixel values in the image formation, the observed image pixel values are close to the neighbor image pixel values. Our idea is to estimate $\mathbf{u}_{est}$, the missing observed image pixel $\mathbf{u}$ by using its neighbor image pixel values $\mathbf{v}$ by, for instance, splines interpolation method [19]. The regularization functional $\widetilde{\mathcal{R}}(\mathbf{u})$ is defined to be

$$\widetilde{\mathcal{R}}(\mathbf{u}) \equiv \frac{1}{2}\|\mathbf{u} - \mathbf{u}_{est}\|_2^2.$$

Similarly, an alternating minimization algorithm is developed for solving the joint minimization model (12).

Alternating Minimization Algorithm

Given $\mathbf{u}^0$: iterating $k = 1,2,\ldots$, until convergence

▲ Step i) Determine $\mathbf{f}_k = \arg\min_{\mathbf{f}} J_{IL}(\mathbf{f},\mathbf{u}_{k-1})$ by solving the corresponding Euler-Lagrange equation:

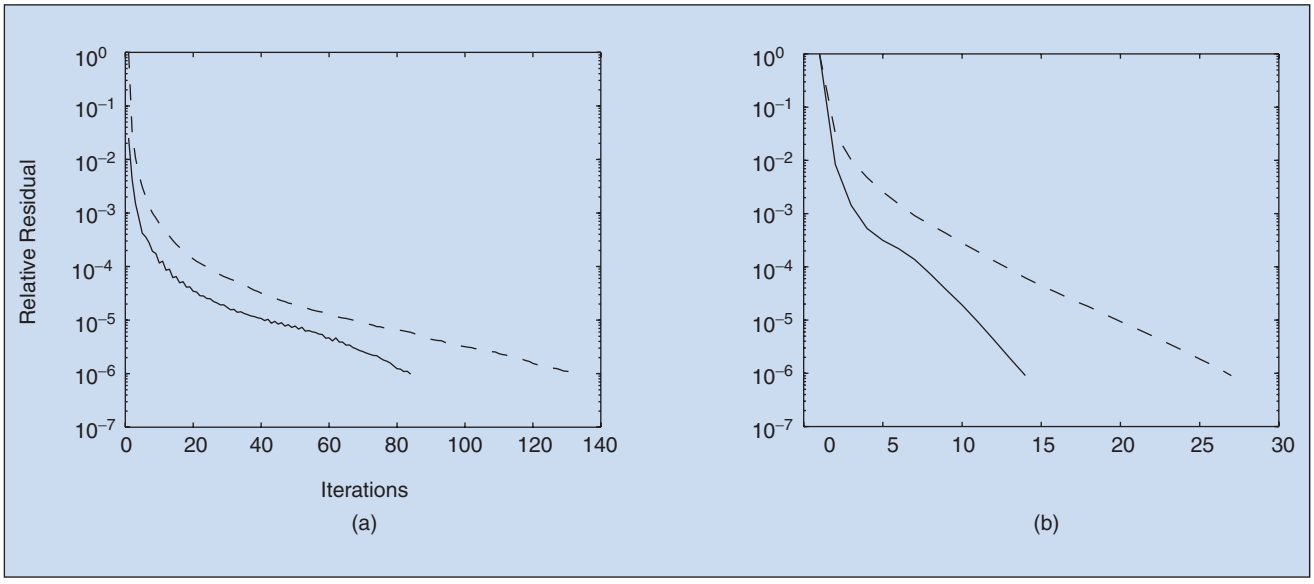
$$\left(H^T(\mathbf{e})H(\mathbf{e}) + \alpha_1 R\right)\mathbf{f}_k = H^T(\mathbf{e})(\mathbf{u}_{k-1} + \mathbf{v}).$$

▲ Step ii) Solve $\mathbf{u}_k = \arg\min_{\mathbf{u}} J(\mathbf{f}_k,\mathbf{u})$ by solving the corresponding Euler-Lagrange equation:

$$(1+\alpha_2)\mathbf{u}_k - \sum_{(l_1 l_2)\in \text{missing sensors}} (I - D_{l_1 l_2}) \times H(\mathbf{e})\mathbf{f}^k - \mathbf{v} - \alpha_2 \mathbf{u}_{\text{est}} = 0.$$

For Step i), the linear system can again be solved by the preconditioned conjugate gradient method with the cosine transform preconditioner. Therefore the linear system can be solved very efficiently. For Step ii), the vector $\mathbf{u}_k$ can be computed by using the matrix-vector multiplication. Therefore the proposed algorithm for the joint minimization model is more efficient than the direct application of the cosine transform preconditioner for model (2). Moreover, we see in Figures 10 and 11 that the quality of reconstructed images using the joint minimization model in (12) is better than those using the model in (2).

The convergence of the above alternating minimization algorithm can be analyzed by using the framework of fixed point theory. We just combine the two steps of the alternating minimizing algorithm and derive



▲ *9. Convergence behavior of the original system (- - -) and the cosine transform preconditioned system (—) when SNRs are (a) 50 dB and (b) 30 dB for missing two LR images.*

$$\mathbf{f}_k = \frac{1}{1+\alpha_2}[H^T(\mathbf{e})H(\mathbf{e})+\alpha_1 R]^{-1} H^T(\mathbf{e}) \times \sum_{(l_1 l_2)\in \text{missing sensors}} (I - D_{l_1 l_2})H(\mathbf{e})\mathbf{f}_{k-1} + \left[H^T(\mathbf{e})H(\mathbf{e})+\alpha_1 R\right]^{-1} \times H^T(\mathbf{e})\left(\frac{\alpha_2}{1+\alpha_2}\mathbf{u}_{\text{est}} + \frac{1}{1+\alpha_2}\mathbf{v}\right).$$

In [37], it has been shown that the spectral radius of the matrix

$$\frac{1}{1+\alpha_2}\left[H^T(\mathbf{e})H(\mathbf{e})+\alpha_1 R\right]^{-1} \times H^T(\mathbf{e}) \sum_{(l_1 l_2)\in \text{missing sensors}} (I - D_{l_1 l_2})H(\mathbf{e})$$

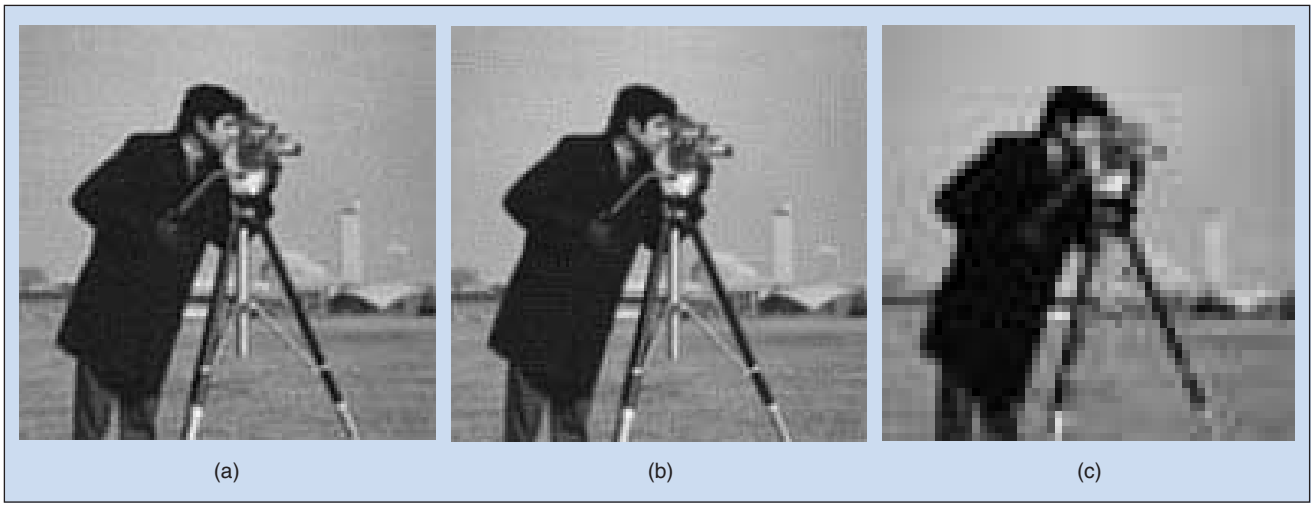
If we consider edge-preserving regularization models, can we design an efficient iterative method for solving the corresponding minimization problem?

is less than one, and therefore the alternating minimization algorithm converges globally to a minimizer for any given initial guess.

Edge-Preserving Regularization

The H_1-norm regularization tends to attenuate the high frequency information of the HR image. Therefore,

▲ *10. (a) Original and (b), (c) two LR images.*



▲ *11. Reconstructed images using (a) the model in (2) [relative error = 0.0614], (b) the joint minimization model with the H^1-norm regularization [relative error = 0.0574], and (c) the joint minimization model with the total variation regularization [relative error = 0.0931].*

The key problem of simultaneously estimating an unknown blur from the observed sequence and attaining SR of a region of interest remains to be fully tackled.

edge-preserving regularization is considered. In Figures 11 and 12, we see that the quality of reconstructed images using the edge-preserving regularization is better than those using the H_1-norm regularization.

In this section, we focus on the regularization term Φ of the form

$$\Phi(\mathbf{f}) = \sum_{i=1}^{r} \phi(\mathbf{d}_i^T \mathbf{f}), \tag{13}$$

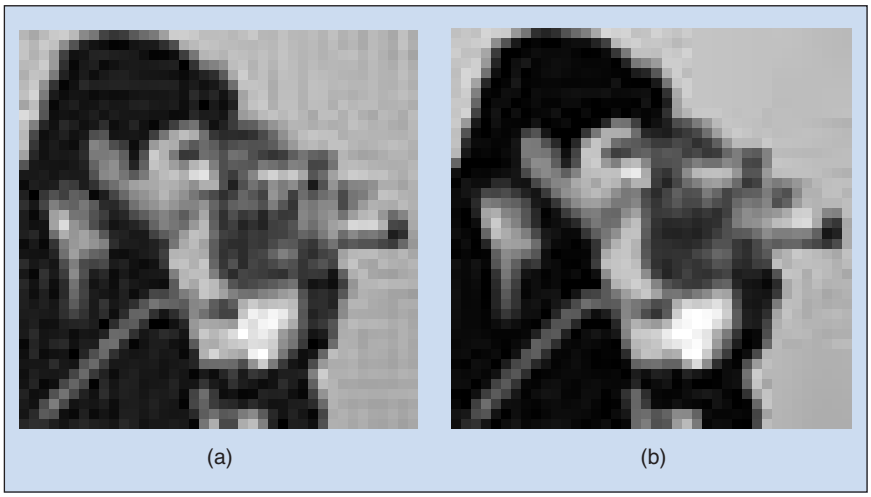
where $\phi: \mathbb{R} \to \mathbb{R}$ is a potential function and $\mathbf{d}_i^T$, for $i = 1, \ldots, r$, are linear operators. Typically, the elements of the sequence $\{\mathbf{d}_i^T \mathbf{f}\}$ are either first- or second-order differences between neighboring pixels and $\mathbf{r}$ is the number of these differences computed from the image. Assume that ϕ is smooth, convex, and *edge preserving*, i.e., $\phi(t) < t^2$ as $|t| \to \infty$. Examples of such functions are [2], [4], [9], [11]:

$$\phi(t) = |t|^{\beta}, 1 < \beta \le 2,$$

$$\phi(t) = \sqrt{\beta + t^2},$$

$$\phi(t) = \begin{cases} t^2/2 & \text{if } |t| \le \beta, \\ \alpha|\tau| - \beta^2/2 & \text{if } |t| > \beta. \end{cases}$$

However, the resulting minimizers of such objective functions are nonlinear with respect to $\mathbf{f}$ and computation intensive.



(a) (b)

▲ 12. (a) H_1-norm regularization and (b) the edge-preserving regularization.

To cope with numerical slowness, half-quadratic (HQ) reformulation has been used [12], [13]. The idea is to construct an augmented cost function that involves an auxiliary variable. In the HR image reconstruction, the augmented cost-function is given by

$$J(\mathbf{f}, \mathbf{s}) = \|H(\mathbf{e})\mathbf{f} - \mathbf{g}\|_2^2 + \frac{\alpha}{2}\|L\mathbf{f} - \mathbf{s}\|_2^2 + \beta \sum_i \psi(s_i). \tag{14}$$

This formulation is considered under the condition that the function

$$t \to t^2/2 - \phi(t)$$

is convex, continuous, and finite for every $t \in \mathbb{R}$ and the following expressions are equivalent:

$$\phi(t) = \inf_{s \in R}\left\{\psi(s) + (t-s)^2/2\right\},$$

$$\psi(s) = \sup_{t \in R}\left\{\phi(t) - (t-s)^2/2\right\}.$$

The augmented cost function is minimized using an alternating minimization scheme. Let the solution obtained at iteration $(k-1)$ read $(\mathbf{f}^{(k-1)}, \mathbf{s}^{(k-1)})$. At the next iteration k one calculates

$$\mathbf{s}^{(k)} \text{ such that } \quad J\left(\mathbf{f}^{(k-1)}, \mathbf{s}^{(k)}\right) \le J\left(\mathbf{f}^{(k-1)}, \mathbf{s}\right), \quad \forall \mathbf{s} \in \mathbb{R}^r,$$

$$\mathbf{f}^{(k)} \text{ such that } \quad J\left(\mathbf{f}^{(k)}, \mathbf{s}^{(k)}\right) \le J\left(\mathbf{f}, \mathbf{s}^{(k)}\right), \quad \forall \mathbf{f} \in \mathbb{R}^{n^2}.$$

These minimizations give rise to two minimizer mappings, $\mathbf{f} \to [\xi_s(\mathbf{d}_1^T \mathbf{f}), \ldots, \xi_s(\mathbf{d}_r^T \mathbf{f})]^T$ with $\sigma: \mathbb{R} \to \mathbb{R}$, and $s \to \chi(s)$ with $\chi: R^r \to R^{n^2}$. The alternate minimization thus reads

$$s_i^{(k)} = \sigma\left(\mathbf{d}_i^T \mathbf{f}^{(k-1)}\right), \ \forall i = 1, \ldots, r,$$

$$\mathbf{f}^{(k)} = \left(H^T(\mathbf{e})H(\mathbf{e}) + \alpha L^T L\right)^{-1}\left(2H^T(\mathbf{e})\mathbf{g} + \beta L^T \mathbf{s}\right)$$

where the minimizer function σ reads [11], [42]:

$$\sigma(t) = t - \phi'(t).$$

The above linear system can be solved efficiently by the cosine transform-based conjugate gradient method (see [41]). Although the intuition that HQ regularization does indeed increase the speed of the minimization of regularized cost functions of the form (14), the critical question has not been considered in a theoretical way. The convergence of the edge-preserving regularization will be further studied in depth.

Future Research

It is usually not possible at the outset to achieve the desired resolution because of technology and cost constraints. For example, the technology of a CCD is limited by factors like physical dimension, shot noise, and parasitic effects. In applications like astronomical imaging, the reduced size and weight of cameras in a spaceship or satellite affect its quality. The need for tradeoff between size, weight, and quality of the CCD array necessitates the design of SR algorithms to obtain the desired HR image of a common region of interest in the frames of a video sequence without modifying the physical characteristics of the CCD array. Imperfect optics, finite detector arrays, and finite individual detector sizes all contribute to a variety of degradation processes to which an image acquisition system is susceptible. Dramatic progress has been documented during the last decade in the area of HR image processing that encompasses the stages of image registration or camera motion parameter estimation, deblurring, noise reduction (filtering), and interpolation.

Considerable research activity is being witnessed in areas pertaining to the construction of a panoramic mosaic followed by the attainment of spatial resolution increase of regions of interest in the mosaic. In video, the user gets the sequence of images containing both spatial and temporal information. The number of pixels in each frame is fixed to a known resolution. To generate one big snapshot of the scene that covers all desired areas (panoramic mosaic), one needs to adjust the camera in various ways to capture the effects of zooming, panning, tilting, etc., that may be required for capturing the entire scene. Usually, the resulting picture will suffer from undersampling or LR effect because the whole scene has to be represented by the limited number of pixels. Therefore, the bigger the scene, the lower will be the resolution. On the other hand, higher resolution of regions of interest in the scene can be obtained if the camera is zoomed to those specific regions. Therefore the tradeoff between the size of the scene and its resolution has to be addressed. In the future, more research is needed on how to take advantage of the intraframe spatial information along with interframe temporal information to create the HR panoramic image and then attain SR of regions of interest in the mosaic.

The key problem of simultaneously estimating an unknown blur from the observed sequence and attaining SR of a region of interest remains to be fully tackled. Blind deconvolution refers to the problem of restoring the original image from a degraded observation and incomplete blur information. Existing blind deconvolution techniques can be categorized into two main classes. One approach treats blur identification and image super resolution separately. The other approach implements the two subtasks simultaneously. Both approaches are adaptable from the single image case to the multiframe situation of interest here. Initial results in this area of blind robust super resolution have been reported recently [26]. Recently, Nguyen et al. [40] considered the blind deconvolution problem in the multiframe case. Their algorithm can handle only the PSF, which is modeled by one parameter. This restriction is very limited in many applications.

Acknowledgments

This research was supported in part by RGC Grants HKU 7132/00P and 7130/02P and by Army Research Office Grant DAAD 19-00-1-0539.

Michael K. Ng received the B.Sc. and M.Phil. degrees in mathematics from the University of Hong Kong in 1990 and 1993, respectively, and the Ph.D. degree in mathematics from the Chinese University of Hong Kong, in 1995. From 1995 to 1997 he was a research fellow at the Australian National University, Canberra. He is currently an associate professor in the Department of Mathematics at the University of Hong Kong. His research interests are in the areas of data mining, operations research, and scientific computing. He was one of the recipients of the Outstanding Young Researcher Award of the University of Hong Kong in 2001.

Nirmal K. Bose received the B. Tech (Hons.), M.S., and Ph.D. degrees in electrical engineering from I.I.T., Kharagpur, India, Cornell University, and Syracuse University, respectively. He was a professor of electrical engineering and of mathematics at the University of Pittsburgh. He joined Pennsylvania State University, University Park, in 1986 as Singer Professor and in 1992 was named the HRB-Systems Professor of electrical engineering. He is the author of several recognized texts and since is the founding editor-in-chief of the *International Journal on Multidimensional Systems and Signal Processing*. He has served as visiting faculty at several institutions, including the American University of Beirut, Lebanon, and Princeton University, New Jersey. He is a Fellow of the IEEE. His most recent honors include the Invitational Fellowship from the Japan Society for the Promotion of Science in 1999, the Alexander von Humboldt Research Award from Germany in 2000, and the Charles H. Fetter University Endowed Fellowship in Electrical Engineering from 2001-2004.

References

[1] H. Andrew and B. Hunt, *Digital Image Restoration*. Englewood Cliffs, NJ: Prentice-Hall, 1977.

[2] G. Aubert and L. Vese, "A variational method in image recovery," *SIAM J. Numer. Anal.*, vol. 34, pp. 1948-1979, 1997.

[3] M. Banham and A. Katsaggelos, "Digital image restoration," *IEEE Signal Processing Mag.*,vol. 14, pp. 24-41, Mar. 1997.

[4] M. Black and A. Rangarajan, "On the unification of line processes, outlier rejection, and robust statistics with applications to early vision," *Int. J. Comput. Vision*, vol. 19, pp. 57-91, 1996.

[5] K. Boo and N.K. Bose, "Two-dimensional model-based power spectrum estimation for nonextendible correlation bisequences," *Circuits Syst. Signal Process.*, vol. 16, no. 2, pp. 141-163, 1997.

[6] N.K. Bose and K. Boo, "Asymptotic eigenvalue distribution of block-Toeplitz matrices," *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 858-861, 1998.

[7] N.K. Bose, S. Lertrattanapanich, and J. Koo, "Advances in superresolution using the L-curve," in *Proc. Int. Symp. Circuits and Systems*, Sydney, Australia, May 2001, pp. 433-436.

[8] N.K. Bose and K. Boo, "High-resolution image reconstruction with multisensors," *Int. J. Imaging Syst. Technol.*, vol. 9, pp. 294-304, 1998.

[9] C. Bouman and K. Sauer, "A generalized Gaussian image model for edge-preserving MAP estimation," *IEEE Trans. Image Processing*, vol. 2, pp. 296-310, July 1993.

[10] R. Chan, T. Chan, M. Ng, W. Tang, and C. Wong, "Preconditioned iterative methods for high-resolution image reconstruction with multisensors," in *Proc. SPIE Symp. Advanced Signal Processing: Algorithms, Architectures, and Implementations,* vol. 3461, San Diego, CA, July, 1998, pp. 348-357.

[11] P. Charbonnier, L. Blanc-Feraud, G. Aubert, and M. Barlaud, "Deterministic edge-preserving regularization in computer imaging," *IEEE Trans. Image Processing*, vol. 6, pp. 298-311, Feb. 1997.

[12] D. Geman and G. Reynolds, "Constrained restoration and recovery of discontinuities," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 14, pp. 367-383, Mar. 1992.

[13] D. Geman and C. Yang, "Nonlinear image recovery with half-quadratic regularization," *IEEE Trans. Image Processing*, vol. 4, pp. 932-946, July 1995.

[14] J. Gillete, T. Stadtmiller, and R. Hardie, "Aliasing reduction in staring infrared images using subpixel techniques," *Opt. Eng.*, vol. 34, pp. 3130-3137, Nov. 1995.

[15] G. Golub and C. Van Loan, *Matrix Computations*, 2nd ed. Baltimore, MD: Johns Hopkins Univ. Press, 1989.

[16] R. Gonzalez and R. Woods, *Digital Image Processing*. New York: Addison Wesley, 1992.

[17] P. Hansen and D. O'Leary, "The use of the L-curve in the regularzation of discrete ill-posed problems," *SIAM J. Sci. Comput.,* vol. 14, pp. 1487-1503, 1993.

[18] M. Hestenes and E. Steifel, "Methods of conjugate gradient for solving linear systems," *J. Res. Nat. Bureau Stand.,* vol. 49, pp. 409-436, 1952.

[19] H. Hou and H. Andrews, "Cubic splines for image interpolation and diagonal filtering," *IEEE Trans. Accoust Speech Signal Processing*, vol. 26, pp. 508-517, Dec. 1978.

[20] B. Hunt and O. Kubler, "Karhunen-Loeve multispectral image restoration, part I: Theory," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. 32, pp. 592-600, June 1984.

[21] G. Jacquemod, C. Odet, and R. Goutte, "Image resolution enhancement using subpixel camera displacement," *Signal Processing*, vol. 26, pp. 139-146, 1992.

[22] T. Komatsu, K. Aizawa, T. Igarashi, and T. Saito, "Signal-processing based method for acquiring very high resolution images with multiple cameras and its theoretical analysis," *Proc. Inst. Elec. Eng.* vol. 140, no. 3, pt. I, pp. 19-25, 1993.

[23] J. Koo and N.K. Bose, "Spatial restoration with reduced boundary error," in *Proc. Mathematical Theory of Networks and Systems (MTNS)*, Univ. of Notre Dame, South Bend, IN., Aug. 12-16, 2002 [Online]. Available: http://www.ndu.edu/~mtns/talksalph.html

[24] R. Lagendijk and J. Biemond, *Iterative Identification and Restoration of Images*. Norwell, MA: Kluwer, 1991.

[25] S. Lertrattanapanich and N.K. Bose, "Latest results on high-resolution reconstruction from video sequences," Inst. Electronic, Information and Communication Eng., Japan, Tech. Rep. IEICE, DSP99-140, Dec. 1999, pp. 59-65.

[26] S. Lertrattanapanich and N.K. Bose, "High resolution image formation from low resolution frames using Delaunay triangulation," *IEEE Trans. Image Processing,* vol. 11, no. 12, pp. 1427-1441, Dec. 2002.

[27] D. Luenberger, *Linear and Nonlinear Programming*, 2nd ed. Reading, MA: Addison-Wesley,1984.

[28] F. Luk and D. Vandevoorde, "Reducing boundary distortion in image restoration," in *Proc. SPIE 2296, Advanced Signal Processing Algorithms, Architectures and Implementations VI*, 1994, pp. 554-565.

[29] S. Mann and R. Picard, "Video Orbits of the projective group: A simple approach to featureless estimation of parameters," *IEEE Trans. Image Processing,* vol. 6, pp. 1281-1295, Sept. 1997.

[30] M. Ng, "An efficient parallel algorithm for high-resolution color image reconstruction," *Proc. the Seventh Int. Conf. Parallel and Distributed Systems: Workshops*, Iwate, Japan, 4-7 July, 2000, pp. 547-552,

[31] M. Ng, N.K. Bose, and J. Koo, "Constrained total least squares computations for high resolution image reconstruction with multisensors," *Int. J. Imaging Syst. Technol.,* vol. 12, pp. 35-42, 2000.

[32] M. Ng and N.K. Bose, "Analysis of displacement errors in high-resolution image reconstruction with multisensors," *IEEE Trans. Circuits Syst. I*, vol. 49, pp. 806-813, June 2002.

[33] M. Ng and N.K. Bose, "Fast color image restoration with multisensors," *Int. J. Imaging Syst. Technol.*, to be published.

[34] M. Ng, N.K. Bose, and J. Koo, "Constrained total least squares for color image reconstruction," in *Total Least Squares and Errors-in-Variables Modelling III: Analysis, Algorithms and Applications*, S. Huffel and P. Lemmerling, Eds. Norwell, MA: Kluwer, 2002, pp. 365-374.

[35] M. Ng, R. Chan, and W. Tang, "A fast algorithm for deblurring models with Neumann boundary conditions," *SIAM J. Sci. Comput.,* vol. 21, pp. 851-866, 1999.

[36] M. Ng, R. Chan, T. Chan, and A. Yip, "Cosine transform preconditioners for high resolution image reconstruction," *Linear Algebra Applicat.,* vol. 316, pp. 89-104, 2000.

[37] M. Ng, W. Ching, K. Sze, and A. Yau, "Super-resolution image reconstruction using multisensors," *Numerical Linear Algebra with Applications*, to be published.

[38] M. Ng and K. Sze, "Preconditioned iterative methods for superresolution image reconstruction with multisensors," *SPIE, Symp. Advanced Signal Processing: Algorithms, Architectures and Implementations,* vol. 4116, San Diego CA, 2000, pp. 396-405.

[39] M. Ng and A. Yip, "A fast MAP algorithm for high-resolution image reconstruction with multisensors," *Multidimensional Syst. Signal Process.,* vol. 12, no. 2, pp. 143-164, 2001.

[40] N. Nguyen, P. Milanfar, and G. Golub, "Efficient generalized cross-validation with applications to parametric image restoratioin and resolution enhancement," *IEEE Trans. Image Processing*, vol. 10, pp. 1299-1308, Sept. 2001.

[41] M. Nikolova and M. Ng, "Fast image reconstruction algorithms combining half-quadratic regularization and preconditioning," in *Proc. IEEE Int. Conf. Image Processing*, 2001, vol. I, pp. 277-280.

[42] M. Nikolova and M. Ng, "Comparison of the main forms of half-quadratic regularization," in *Proc. IEEE Int. Conf. Image Processing,* Rochester, NY, Sept. 2002, vol. I, pp. 349-352.

[43] S. Van Huffel and J. Vandewalle, *The Total Least Squares Problem: Computational Aspects and Analysis*. Philadelphia, PA: SIAM, 1991.