

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/44589937>

Efficient Fourier-Wavelet Super-Resolution

Article in *IEEE Transactions on Image Processing* · May 2010

DOI: 10.1109/TIP.2010.2050107 · Source: PubMed

CITATIONS

62

READS

155

4 authors, including:



Cynthia Toth
Duke University

370 PUBLICATIONS 11,167 CITATIONS

[SEE PROFILE](#)



Joseph Y Lo
Duke University Medical Center

265 PUBLICATIONS 4,517 CITATIONS

[SEE PROFILE](#)



Sina Farsiu
Duke University Medical Center

214 PUBLICATIONS 9,331 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Stereoviewing and Breast Tomosynthesis [View project](#)



Comparison of AMD Treatment Trials [View project](#)

Efficient Fourier-Wavelet Super-Resolution

M. Dirk Robinson, Cynthia A. Toth, Joseph Y. Lo, and Sina Farsiu

Abstract—Super-resolution (SR) is the process of combining multiple aliased low-quality images to produce a high-resolution high-quality image. Aside from registration and fusion of low-resolution images, a key process in SR is the restoration and denoising of the fused images. We present a novel extension of the combined Fourier-wavelet deconvolution and denoising algorithm ForWarD to the multiframe SR application. Our method first uses a fast Fourier-base multiframe image restoration to produce a sharp, yet noisy estimate of the high-resolution image. Our method then applies a space-variant nonlinear wavelet thresholding that addresses the nonstationarity inherent in resolution-enhanced fused images. We describe a computationally efficient method for implementing this space-variant processing that leverages the efficiency of the fast Fourier transform (FFT) to minimize complexity. Finally, we demonstrate the effectiveness of this algorithm for regular imagery as well as in digital mammography.¹

Index Terms—Digital X-ray imaging, multiframe deblurring, super-resolution (SR), wavelets, denoising.

I. INTRODUCTION

SUPER-RESOLUTION (SR) is the process of combining multiple aliased low-resolution (LR) images to produce a high-resolution high-quality image at a resolution greater than the sampling rate of the detector. SR has received much attention in recent years in the image processing community. We refer the interested reader to [2]–[4] for a broad review of recent algorithmic development in this area.

Aside from registration and fusion of low-resolution images, a key process in SR is the restoration and denoising of fused images. In this paper, we propose an efficient restoration and denoising method that is a novel multiframe extension of the Fourier wavelet regularized deconvolution (ForWarD) algorithm [5], which considers the nonstationarity of the multiframe reconstruction process. The algorithm's efficiency stems from separating the multiframe deconvolution or restoration step from the wavelet-based denoising step allowing us to achieve

nonlinear denoising in a noniterative fashion. Furthermore, we describe how to efficiently implement the algorithm to address the computational complexity associated with the nonstationary noise processes inherent to multiframe reconstruction.

In this paper, we also explore an application of this algorithm to digital mammography. While relatively new, digital mammography is rapidly replacing film-based mammography for the screening and diagnosis of early carcinomas in women. Solid-state detectors have demonstrated improved performance in terms of specificity and sensitivity over film-based imaging for certain groups of women such as those with dense breast tissue, women under the age of fifty, and premenopausal women [6].

Unlike film-based mammography, digital mammography provides the opportunity to directly apply sophisticated digital processing techniques without the need for a secondary film scanning process. An ideal digital mammography system exposes the patient to the minimum amount of radiation required to accomplish the screening task. Digital mammography systems face the same design tradeoff between image resolution, signal-to-noise ratio (SNR), and illumination or radiation exposure level as those found in any digital imaging system. Shrinking the pixel dimension at the detector increases sampling resolution at the expense of dynamic range and SNR. While improved SNR and dynamic range may be obtained by combining multiple images, increasing total radiation beyond the standard dosage is undesirable for the safety of the patient. Alternatively, using large detector pixels improves both the dynamic range and the SNR of the system at the obvious expense of resolution. Digital mammography imaging systems typically choose the highest resolution which supports a minimum required SNR.

To overcome the said quality tradeoffs, we propose digitally combining multiple low-dosage images, each containing spatial shifts. This motion may be the result of patient movement, intentional dithering of the detector, vibration in the imaging system, or small movement of the imaging gantry. In practice, the motion contained in the captured images is a combination of all such sources necessitating accurate registration of the aliased low-resolution (LR) images.

Applying SR algorithms to digital mammography has two inherent challenges. The captured low-resolution images are typically of size 10 megapixels and larger. Thus, algorithmic efficiency is very important due to the sheer size of the reconstructed images, which could range from 40 to 160 megapixels depending on the resolution enhancement factor. Also, to minimize total radiation exposure, we must use lower than normal dosages of X-ray exposure for capturing each frame. Therefore, the captured data has extremely low peak SNR (PSNR). For example, Fig. 1 compares a high dosage X-ray image (computed

Manuscript received January 19, 2009; revised August 31, 2009; accepted February 03, 2010. Date of publication May 10, 2010; date of current version September 17, 2010. This work was supported in part by Siemens Healthcare. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Arun Ross.

M. D. Robinson is with Ricoh Innovations, Menlo Park, CA 94025 USA (e-mail: dirkr@rii.ricoh.com).

C. A. Toth and S. Farsiu are with the Departments of Ophthalmology and Biomedical Engineering, Duke University, Durham, NC 27710 USA (e-mail: cynthia.toth@duke.edu; sina.farsiu@duke.edu).

J. Y. Lo is with the Departments of Radiology and Biomedical Engineering, Duke University, Durham, NC 27710 USA (e-mail: joseph.lo@duke.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIP.2010.2050107

¹Preliminary results of this work were presented at ICIP, October 2008 [1].

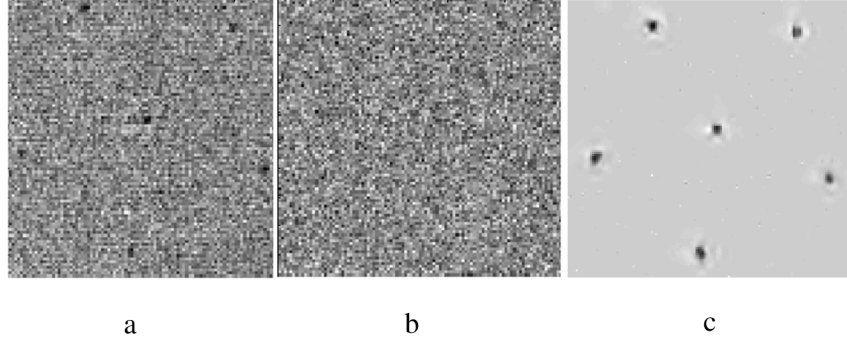


Fig. 1. Mammogram X-ray images from a phantom breast containing a pentagram-shaped set of micro-calcification cluster. (a) High dosage at 226mAs, faintly showing the nodules (PSNR ≈ 13 dB). (b) Extremely low-dosage at 11.3 mAs used in the proposed multiframe reconstruction scheme (PSNR ≈ 3 dB), where the nodules are almost undetectable. The total dosage of using 15 of these frames ($15 \times 11.3 = 170$ mAs) is still less than the high dosage. (c) Restoration combining the 15 low-dosage frames, clearly demonstrating the pentagram-shaped set of micro-calcification cluster.

PSNR² ≈ 13 dB) with the very low exposure images (computed SNR ≈ 3 dB) used in our multiframe scheme. We demonstrate that the effectiveness of the two-stage restoration and denoising algorithm allows us to provide high-resolution, high contrast, and low noise images at very low radiation dosages. Furthermore, our results suggest improved detection rates of texture resembling small calcification in breast tissue. These results suggest new tradeoffs in designing digital mammogram systems.

In Section II, we describe the forward imaging model and the problem of SR. In Section III, we describe the multiframe variant of the ForWarD algorithm we employ to restore and denoise the reconstructed images. Section IV presents experimental results using this new approach and Section V outlines some future directions of this work.

II. BACKGROUND: SR AND EFFICIENT DEBLURRING

In this section, we establish the background required for the development of our efficient multiframe SR algorithm in Section III. We introduce our imaging model and since our method is a combination of the multiframe SR and ForWarD deblurring algorithms, we also review these two concepts. Moreover, we review a novel multiframe motion estimation algorithm [7] that is used to produce the results in Section IV. We believe that due to some algorithmic similarities, this brief review of the motion estimation technique facilitates the study of the method described in Section III.

A. Imaging Model

The captured LR X-ray images are often very large and may contain complicated relative motions due to patient motion. However, following several other popular SR methods [8]–[10], we consider the translational (or pure rotational) motion models. To better justify and extend the application of this model, in [11], we introduced a novel joint motion estimation and SR approach in a tile-based fashion. That is, instead of considering a global translational motion model, we assume that each LR image is made of a set of small sized

tiles (blocks). These tiles move independently in a translational (or pure rotational) model. When using tile-based processing, the relative shifts between data sets is better approximated by the translational (or rotational) motion models. Furthermore, applying the algorithm to small tiles reduces the memory requirements of the multiframe reconstruction algorithm. The motion estimation process involves identifying image tiles from different LR images corresponding to a particular region of interest. In [11], we introduced a joint estimation technique, in which matching blocks of different LR frames are optimally detected and registered in a multiframe joint estimation process. To simplify the notations, without the loss of generality, all formulas used in this paper correspond to the reconstruction of a single HR tile (a full image is reconstructed by stitching a set of such HR tiles).

We denote the raster scanned version for each of the $K+1$ LR image tiles by the vector \mathbf{y}_k . These noisy LR input image tiles are blurry, translated, and downsampled versions of an unknown high-resolution image tile denoted by \mathbf{x} . The forward model relating these captured image tiles to the unknown high-resolution tile is given by

$$\mathbf{y}_k = \mathbf{DHS}(\mathbf{v}_k)\mathbf{x} + \mathbf{e}_k \quad (1)$$

in which the vector \mathbf{y}_k represents $B \times B$ (assumed square without loss of generality) samples of the captured image $y_k(m'_1, m'_2)$, where $m'_1, m'_2 \in [0, B-1]$, are ordered as a $(B)^2 \times 1$ vector. The captured image is undersampled with respect to an unknown high-resolution image $x(m_1, m_2)$, where $m_1, m_2 \in [0, DB-1]$, by a factor of D in each dimension. The vector \mathbf{x} represents samples of the unknown $DB \times DB$ high-resolution image tile $x(m_1, m_2)$ similarly ordered. The matrix \mathbf{H} represents the blurring associated with the imaging system. In X-ray imaging, this blurring is due to the geometry of the radiating illumination as well as the scattering of light in the object material. In each set of tiles, we approximate this effect by a spatially-invariant point spread function (PSF) $h(m_1, m_2)$. Although, the PSF can be different in different tile sets. The warping operator $\mathbf{S}(\mathbf{v}_k)$ of size $(DB)^2 \times (DB)^2$ represents the subpixel spatial shifts between similar tiles in the captured images. The spatial shifting is described by the vector $\mathbf{v}_k = [v_{k1}, v_{k2}]$ for the k th frame. In our model, we

²In this work, the PSNR was computed numerically as $\text{PSNR} = 20 \log_{10}(s/n)$. In experiments on real images, s is the grayscale difference between the minimum and maximum signal regions and n is the noise standard deviation estimated from flat regions. In simulated experiments, n is the RMSE error between the estimated and ground truth image.

assume that these spatial shifts are continuous values in the range of $[-D, D]$. This corresponds to the range of subpixel motions in the captured images. Without loss of generality, we assume that the tile \mathbf{y}_0 defines the coordinate system of the high-resolution image and, hence, we only have to estimate the unknown motion parameters for the remaining K images. The downsampling operator \mathbf{D} of size $B^2 \times (DB)^2$ captures the undersampling of the detector. Finally, \mathbf{e}_k of size $B^2 \times 1$ represents the noise inherent in the analog-to-digital conversion. For our purposes, we assume this noise to be uncorrelated zero-mean noise with standard deviation σ .

B. Classic Maximum A Posteriori SR Reconstruction

The general problem of SR is to combine $K + 1$ captured LR images and estimate the high-resolution image \mathbf{x} . An important category of solutions to the SR problem is the *maximum a posteriori* (MAP) and closely related methodologies, which is discussed in this section. The MAP methods are based on the construction of a cost function (Ω), which is the summation of two distinct terms. One is the data penalty term Ω_d , which measures the closeness of data to the estimates. The other is the regularization term Ω_p , which represents the prior information about the unknown high-resolution (HR) image (\mathbf{x}).

Early MAP-based SR methods assumed that the motion vectors were accurately estimated in a separate process and the noise model was Gaussian [12], [13], which justifies the application of quadratic data penalty terms. As for the regularization term, these techniques most frequently employed the quadratic Tikhonov style regularization despite its tendency to reduce edge contrast. The resulting cost function is in the form of

$$\Omega(\mathbf{x}) = \sum_{k=0}^K \|\mathbf{y}_k - \mathbf{DHS}(\mathbf{v}_k)\mathbf{x}\|^2 + \lambda \mathbf{x}^T \mathbf{C}_\mathbf{x}^{-1} \mathbf{x} \quad (2)$$

where $\mathbf{C}_\mathbf{x}^{-1}$ is often a spatial high-pass operator and λ is the weighting scalar. When $\mathbf{C}_\mathbf{x}$ is the exact covariance of the unknown HR image, then this cost function produces the ideal Wiener filter estimate of the unknown image. This MAP functional has the advantage of being quadratic, which means that the estimate image is a linear function of the input measurements and is, thus, easy to compute.

Through the years, application of more advanced prior functions Ω_p such as Adaptive Kernel regression [14] which generalizes popular priors such as Tikhonov and Bilateral Total-Variation (B-TV) [9], have produced higher quality estimates. For example, the B-TV cost function is defined as

$$\Omega_{\text{B-TV}}(\mathbf{x}) = \sum_{k=0}^K \|\mathbf{y}_k - \mathbf{DHS}(\mathbf{v}_k)\mathbf{x}\|_2^2 + \lambda \sum_{t_1, t_2=-L}^L \varrho^{|t_1|+|t_2|} \|\mathbf{x} - \mathbf{S}(\mathbf{t})\mathbf{x}\|_1 \quad (3)$$

where $\mathbf{t} = [t_1, t_2]$ is a set of integer pixel shifts and $0 < \varrho \leq 1$ is a constant [9]. The parameter L defines the size of the corresponding Bilateral filter kernel. The Bilateral filter and its parameters are extensively discussed in [15], [9].

These advanced regularization functions are not quadratic (nonlinear estimators) and, hence, require more computationally-complex iterative minimization strategies. Such nonquadratic functionals can, however, preserve many important features of images such as edges. Also, MAP-based robust SR techniques (e.g., [9], [16], and [17]) are able to reduce the effect of outliers such as motion estimation error.

Practical tests show that using a separate motion estimation process, specially in low-SNR cases, is suboptimal. Therefore, the critical issue of joint SR and motion estimation problem has been the topic of several papers (e.g., [18]–[22]). A simplified MAP formulation of this problem has the form

$$\Omega(\mathbf{x}, \underline{\mathbf{v}}) = \sum_{k=0}^K \|\mathbf{y}_k - \mathbf{DHS}(\mathbf{v}_k)\mathbf{x}\|^2 + \lambda \mathbf{x}^T \mathbf{C}_\mathbf{x}^{-1} \mathbf{x} \quad (4)$$

where $\underline{\mathbf{v}} = [\mathbf{v}_1, \dots, \mathbf{v}_K]^T$. Note that, additional priors on motion vector distribution may also be added to the previously mentioned cost function [18]. The previously mentioned approaches are commonly solved in an iterative fashion and are relatively computationally expensive. While the joint estimation techniques are generally computationally more complex than robust SR solutions, they are more effective when the number of LR images is small or when the motion of most LR frames are estimated erroneously. As noted in [20], it is only by jointly estimating the unknown motion vectors and the aliasing free image that estimators can avoid the bias associated with registering images containing aliasing artifacts.

C. Problem of Joint Motion and Image Estimation

In this subsection, we briefly review an alternative approach for estimating the image shifts between aliased images using the variable-projection principal [23], which we described in detail in our recent publication [7]. While motion-estimation is not the focus of this paper, study of this technique provides intuition and simplifies the material described in later sections of this paper.

Considering the PSF and motion assumptions in Section II-A, we may reverse the order of the shifting and blur operators in (1) [8] and rewrite the imaging model as

$$\mathbf{y}_k = \mathbf{DS}(\mathbf{v}_k)\mathbf{H}\mathbf{x} + \mathbf{e}_k = \mathbf{DS}(\mathbf{v}_k)\mathbf{z} + \mathbf{e}_k \quad (5)$$

where $\mathbf{z} = \mathbf{H}\mathbf{x}$ is the unknown HR blurry image. The optimization process will then be formulated as

$$\Omega(\mathbf{z}, \underline{\mathbf{v}}) = \sum_{k=0}^K \|\mathbf{y}_k - \mathbf{DS}(\mathbf{v}_k)\mathbf{z}\|_2^2 + \lambda \mathbf{z}^T \mathbf{C}_\mathbf{z}^{-1} \mathbf{z} \quad (6)$$

where $\mathbf{C}_\mathbf{z}$ is the covariance matrix of the unknown signal \mathbf{z} , which is typically assumed to be stationary. A typical solution to the previously mentioned problem is the cyclic coordinate-descent method [18], in which in each iteration one unknown variable is updated based on the estimate of the other unknown variable in the previous iteration.

D. Efficient Joint Estimation Using Variable Projections

Noting that (6) is known in numerical analysis literature as the Separable Nonlinear Least Squares problem [23], in our Variable-Projection technique, we momentarily assume that the non-

linear parameters (motion-vectors) are known. Consequently, the estimate of the set of linear parameters (\mathbf{z}) is computed as

$$\hat{\mathbf{z}} = (\mathbf{Q}(\underline{\mathbf{y}}) + \lambda \mathbf{C}_{\mathbf{z}}^{-1})^{-1} \mathbf{g}(\underline{\mathbf{y}}) \quad (7)$$

where

$$\mathbf{Q}(\underline{\mathbf{y}}) = \frac{1}{\sigma^2} \sum_{k=0}^K \mathbf{S}^T(\mathbf{v}_k) \mathbf{D}^T \mathbf{D} \mathbf{S}(\mathbf{v}_k) \quad (8)$$

$$\mathbf{g}(\underline{\mathbf{y}}) = \frac{1}{\sigma^2} \sum_{k=0}^K \mathbf{S}^T(\mathbf{v}_k) \mathbf{D}^T \mathbf{y}_k. \quad (9)$$

We plug the parametric estimate of the blurry HR image ($\hat{\mathbf{z}}$) into the MAP functional (6) and after some algebraic simplifications, we get a new (maximization) cost function that only relies on the motion-vectors

$$\Omega(\underline{\mathbf{y}}) = \mathbf{g}(\underline{\mathbf{y}})^T (\mathbf{Q}(\underline{\mathbf{y}}) + \lambda \mathbf{C}_{\mathbf{z}}^{-1})^{-1} \mathbf{g}(\underline{\mathbf{y}}). \quad (10)$$

Note that, unlike the cyclic coordinate-descent method, we require no iterations between the sets of parameters since we do not explicitly calculate (7). Indeed, a direct approach to maximize (10) involves inverting a large matrix of size $(DB)^2 \times (DB)^2$ which is computationally challenging for even small image tiles. In [7], we described a series of numerical tricks to speed up the process. One trick is solving the problem in the Fourier domain and taking advantage of the spectral folding phenomenon in aliased images.

To simplify the derivation, we reformulate the image capture model (19) in the Fourier domain. We use the over-script “ \sim ” to denote the Fourier domain representation. For example, the k th LR image is given by $\tilde{y}_k(m'_1, m'_2)$, where the LR spatial frequencies are indexed by m'_1 and m'_2 according to $\omega'_i = e^{(j2\pi m'_i/B)} = f^{m'_i}$. The $f' = e^{(j2\pi/B)}$ term is the base of the discrete Fourier transform of the observed image $y_k(m'_1, m'_2)$. Similarly, the high-resolution spatial frequency coordinates are indexed by m_1 and m_2 , where $\omega_i = e^{(j2\pi m_i/DB)} = f^{m_i}$.

Because the shift \mathbf{S} operator is spatially-invariant, its Fourier representation is a diagonal matrix defined as

$$\tilde{\mathbf{S}}(\mathbf{v}_k) = \text{diag}\{f^{m_1 v_{k1} + m_2 v_{k2}}\}. \quad (11)$$

The downsampling operator \mathbf{D} is *not* spatially-invariant and so its Fourier representation is not diagonal. The downsampling operator is, however, periodic and is conveniently represented by

$$\tilde{\mathbf{D}} = (\mathbf{1}_D^T \otimes \mathbf{I}_B) \otimes (\mathbf{1}_D^T \otimes \mathbf{I}_B) \quad (12)$$

where \otimes represents the Kronecker matrix product, $\mathbf{1}_D$ represents a $[D \times 1]$ vector of all ones, and \mathbf{I}_B represents the identity matrix of B dimension. In the Fourier domain, the vector $\tilde{\mathbf{y}}_k$ corresponds to samples of the spectrum of the captured image $\tilde{y}_k(m'_1, m'_2)$.

The form of matrix $\tilde{\mathbf{D}}$ justifies a decoupling of the observed spatial frequency components in the Fourier domain. In other words, we can consider each LR spatial frequency component

(indexed by m'_1 and m'_2) as an independent observation model given by

$$\tilde{y}_k(m'_1, m'_2) = \mathbf{1}_{D^2}^T \tilde{\mathbf{S}}_{m'_1, m'_2}(\mathbf{v}_k) \tilde{\mathbf{z}}_{m'_1, m'_2} + \tilde{\mathbf{e}}_k(m'_1, m'_2) \quad (13)$$

where the $D^2 \times D^2$ matrices are constructed as

$$\tilde{\mathbf{S}}_{m'_1, m'_2}(\mathbf{v}_k) = \text{diag} \left\{ f^{m'_1 v_{k1} + m'_2 v_{k2}}, f^{(m'_1+B)v_{k1} + m'_2 v_{k2}}, \dots, f^{(m'_1+(D-1)B)v_{k1} + (m'_2+(D-1)B)v_{k2}} \right\} \quad (14)$$

and the vectors are constructed according to

$$\tilde{\mathbf{z}}_{m'_1, m'_2} = [\tilde{z}(m'_1, m'_2), \tilde{z}(m'_1 + B, m'_2), \dots, \tilde{z}(m'_1 + (D-1)B, m'_2 + (D-1)B)]^T.$$

A single spatial frequency in the captured image is a function of the original signal content and summation of the $D^2 - 1$ aliased spectral components. This demonstrates that the reconstruction can be applied to each collection of aliased spectral components independently. Thus, a singlet set of D^2 high-resolution spatial frequency components $\tilde{\mathbf{z}}_{m'_1, m'_2}$ can be estimated as

$$\hat{\tilde{\mathbf{z}}}_{m'_1, m'_2} = (\tilde{\mathbf{Q}}_{0,0}(\underline{\mathbf{y}}) + \lambda \tilde{\mathbf{C}}_{\tilde{\mathbf{z}}, m'_1, m'_2}^{-1})^{-1} \tilde{\mathbf{g}}_{m'_1, m'_2}(\underline{\mathbf{y}}) \quad (15)$$

where

$$\tilde{\mathbf{Q}}_{0,0}(\underline{\mathbf{y}}) = \frac{1}{\sigma^2} \sum_{k=0}^K \tilde{\mathbf{S}}_{0,0}^T(\mathbf{v}_k) \mathbf{1}_{D^2} \mathbf{1}_{D^2}^T \tilde{\mathbf{S}}_{0,0}(\mathbf{v}_k) \quad (16)$$

$$\tilde{\mathbf{C}}_{\tilde{\mathbf{z}}, m'_1, m'_2} = \text{diag}\{\tilde{Z}(m'_1, m'_2), \tilde{Z}(m'_1 + B, m'_2), \dots, \tilde{Z}(m'_1 + (D-1)B, m'_2 + (D-1)B)\}$$

$$\tilde{\mathbf{g}}_{m'_1, m'_2}(\underline{\mathbf{y}}) = \frac{1}{\sigma^2} \sum_{k=0}^K \tilde{\mathbf{S}}_{m'_1, m'_2}^T(\mathbf{v}_k) \mathbf{1}_{D^2} \tilde{y}_k(m'_1, m'_2) \quad (17)$$

in which $\tilde{Z}(m_1, m_2)$ are samples of the signal's power spectral density (PSD) function.

Finally, the motion estimation function of (10) simplifies to

$$\Omega(\underline{\mathbf{y}}) = \sum_{m'_1, m'_2} \tilde{\mathbf{g}}_{m'_1, m'_2}(\underline{\mathbf{y}})^T \times (\tilde{\mathbf{Q}}_{0,0}(\underline{\mathbf{y}}) + \lambda \tilde{\mathbf{C}}_{\tilde{\mathbf{z}}, m'_1, m'_2}^{-1})^{-1} \tilde{\mathbf{g}}_{m'_1, m'_2}(\underline{\mathbf{y}}). \quad (18)$$

Estimating the motion vectors using (18) and the high-resolution image using (15) is significantly faster than using the direct matrix form of (10) and (7). The simplified form requires inverting B^2 small matrices of size $D^2 \times D^2$ as opposed to inverting one very large $D^2 B^2 \times D^2 B^2$ matrix of (7). A set of similar acceleration techniques was described in [7].

So far, ignoring the PSF effects, we have studied a computationally efficient multiframe joint motion estimation and

SR technique in the presence of aliasing. In the next subsection, ignoring the aliasing effects, we study an efficient single frame restoration (deblurring and denoising) method. We combine these two techniques in Section III.

E. Single Frame Deblurring Using ForWarD Algorithm

In this section, we consider the problem of restoring the contrast lost due to the blurring (PSF) inherent to the imaging system. To achieve this, we briefly review the fast two-step ForWarD deblurring algorithm [5]. The ForWarD algorithm combines a Fourier-based regularized deconvolution algorithm with a wavelet-based denoising post processing step. As described in the following, in the first step, a Fourier domain implementation of the Wiener filter reduces blur artifacts while suboptimally magnifying the noise. In the second step, a wavelet-based denoising process reduces the noise artifacts.

The basic ForWarD algorithm addresses the problem of restoring an unknown image signal, which has been blurred by a spatially-invariant operator such as a point spread function (PSF) and corrupted by a stationary noise distribution as in the forward model

$$\mathbf{z}_e = \mathbf{H}\mathbf{x} + \mathbf{e}. \quad (19)$$

The noise is assumed to have a stationary distribution. The first step of the ForWarD algorithm involves inverting the spatially-invariant blurring operator \mathbf{H} using a regularized sharpening filter such as the Wiener filter \mathbf{R} [5], [24]. Such estimation process is formulated as

$$\tilde{\mathbf{x}} = \mathbf{R}\mathbf{z}_e \quad (20)$$

$$\mathbf{R} = (\mathbf{H}^T \mathbf{C}_e^{-1} \mathbf{H} + \lambda \mathbf{C}_x^{-1})^{-1} \mathbf{H}^T \quad (21)$$

where \mathbf{C}_e is the covariance matrix of the stationary measurement noise (typically $\mathbf{C}_e = \sigma^2 \mathbf{I}$) and λ is a weighting factor [25]. For example, in the case of the normal and low-dosage images shown in Fig. 1, the noise standard deviation corresponds to about 20% and 65% of the maximum signal intensity, respectively. Indeed, under certain conditions the Wiener filter can be regarded as a Tikhonov regularization functional [26], which was described in Section II-B. While in the common Wiener filter implementation $\lambda = 1$, in ForWarD algorithm, weighting factor is chosen such that $\lambda < 1$. Such perceivably suboptimal choice of the weighting factor often enhances sharpness at the expense of substantial noise amplification.

The Wiener filter can be implemented efficiently in the Fourier domain using fast Fourier transforms (FFT). Implementation of the Wiener filter in the Fourier domain is founded on the assumption that the blurring operator represented by \mathbf{H} is spatially-invariant and that both the noise and signal are stationary random processes. In this case, the blurring operator can be represented by its transfer function $\tilde{H}(m_1, m_2)$ computed as the Fourier transform of the system's PSF. Also, the signal statistics are represented by a power spectral density function (PSD) $\tilde{\mathcal{X}}(m_1, m_2)$. The application of the Wiener

filter is performed in the Fourier domain producing an estimate of the original signal spectrum $\tilde{\mathbf{x}}$ according to

$$\tilde{x}(m_1, m_2) = \frac{\tilde{H}^*(m_1, m_2) \tilde{z}_e(m_1, m_2)}{|\tilde{H}(m_1, m_2)|^2 + \lambda \frac{\sigma^2}{\tilde{\mathcal{X}}(m_1, m_2)}}. \quad (22)$$

This estimate is converted back into the spatial domain using FFT operators to obtain the estimate $\tilde{x}(m_1, m_2)$.

The second step in the ForWarD process applies a nonlinear denoising filter to the wavelet transform of the sharpened image ($\tilde{\mathbf{x}}$) for the purpose of eliminating the residual noise amplified by the restoration filter ($\tilde{\mathbf{e}} = \mathbf{R}\mathbf{e}$). This nonlinear denoising step is based on adaptive thresholding in the wavelet domain. The redundant (no downsampling) wavelet transform [27] is applied to the sharpened image by convolving the image with a set of scaling and wavelet functions, represented by the matrices Φ_i and Ψ_l , producing a set of scale coefficient images $\tilde{\mathbf{s}}_i$ and wavelet coefficient images $\tilde{\mathbf{w}}_l$. Reducing the wavelet coefficients, or wavelet shrinkage, eliminates the noisy artifacts introduced during the deblurring step according to

$$\hat{w}_l(m_1, m_2) = \tilde{w}_l(m_1, m_2) \frac{1}{1 + \frac{\sigma_l^2(m_1, m_2)}{\gamma_l^2(m_1, m_2)}} \quad (23)$$

where (m_1, m_2) identifies the pixel location, $\sigma_l^2(m_1, m_2)$ represents the noise variance at the l th wavelet space for the (m_1, m_2) pixel, and $\gamma_l^2(m_1, m_2)$ is the wavelet signal power. After shrinking the wavelet coefficients in this manner, the inverse wavelet transform produces the final denoised estimate of the original image $\hat{\mathbf{x}}$. We refer the reader to the original work of [5] for a more complete explanation of this process including visualization of the images at the various steps in the algorithm.

Conceptually similar to the Wiener filter, the wavelet shrinkage reduces the observed wavelet coefficients as a function of the local SNR of the wavelet coefficients. The performance of the wavelet shrinkage depends on the ability to predict the SNR for the wavelet coefficients. Because of the wavelet's spatial locality property, the wavelet shrinkage provides efficient local, signal-dependent denoising.

The value of the noise variance for each wavelet function is given by the covariance of the residual noise $\tilde{\mathbf{e}}$ in the wavelet domain. The covariance matrix of the noise in the wavelet domain is given by $\Psi_l^T \tilde{\mathbf{C}}_e \Psi_l$. Because the Wiener filter is spatially-invariant, the residual noise covariance $\tilde{\mathbf{C}}_e$ is circulant. In other words, the residual noise is stationary due to the spatial invariance of the Wiener filter. Since the redundant wavelet filter [27] is also spatially-invariant, the residual noise power in a particular wavelet space is uniform over the entire image [5]. The residual noise power in the wavelet domain is computed efficiently in the Fourier domain by way of

$$\sigma_l^2 = \sigma^2 \sum_{m_1, m_2=0}^B |\tilde{\Psi}_l(m_1, m_2)|^2 \times \frac{|\tilde{H}(m_1, m_2)|^2}{\left(|\tilde{H}(m_1, m_2)|^2 + \lambda \frac{\sigma^2}{\tilde{\mathcal{X}}(m_1, m_2)}\right)^2}. \quad (24)$$

This must be computed once for each wavelet filter used in the thresholding.

The accuracy of the wavelet thresholding depends on accurate estimates of the local signal statistics in the wavelet domain. In practice, this signal information must be estimated from the captured image data. In [5], the authors propose a simple process for estimating the signal power term $\gamma_l^2(m_1, m_2)$. The authors estimate the wavelet signal power by first very coarsely denoising the sharpened image $\hat{\mathbf{x}}$. They propose a simple hard thresholding wavelet denoising approach to obtain the coarsely denoised image using a *different* set of scaling and wavelet functions Φ'_i and wavelet Ψ'_l than those used for the soft thresholding. We denote the wavelet coefficient images for this different wavelet transform by \mathbf{w}'_l . A hard thresholding applied to the wavelet and scaling coefficients according to

$$\dot{w}'_l(m_1, m_2) = \begin{cases} w'_l(m_1, m_2), & |w'_l(m_1, m_2)| > \kappa \sigma_l \\ 0, & |w'_l(m_1, m_2)| \leq \kappa \sigma_l \end{cases} \quad (25)$$

provides the coarsely denoised image. The κ term is the input threshold ([28] describes the choices for κ). The inverse wavelet transform applied to the hard thresholded wavelet coefficients produces the coarsely denoised image $\dot{\mathbf{x}}'$ (we refer the interested reader to [5] for more information).

Next, the wavelet transform using the *original* wavelet function Φ_i and wavelet Ψ_l is applied to the coarsely denoised image $\dot{\mathbf{x}}'$ to produce the coefficients $\dot{\mathbf{w}}_l$

$$\dot{\mathbf{w}}_l = \Psi_l \dot{\mathbf{x}}'. \quad (26)$$

The spatially-varying standard deviation of the signal's wavelet coefficients is estimated to be the value of the coefficients of the coarsely denoised image, or

$$\gamma_l(m_1, m_2) = \dot{w}_l(m_1, m_2). \quad (27)$$

This estimate of the signal power is used in (23). The ForWarD approach to deconvolution has been applied to several deconvolution problems with success and has been used as a benchmark for evaluating the success of other single frame deconvolution algorithms [14]. In the next section, we propose a multiframe extension of this algorithm, which considers the aliasing and issues inherent to the SR problem.

III. MULTIFRAME IMAGE RESTORATION AND WAVELET DENOISING

The goal of the multiframe SR problem, aside from reducing the aliasing artifacts, is restoring the contrast lost due to the blurring inherent to the imaging system. To achieve this, we derive a novel multiframe variant of the fast two-step ForWarD method [5]. The authors of the original ForWarD algorithm have recently proposed a new version of this algorithm that addresses the multiframe deblurring problem [24] for nonaliased images with stationary noise model. In this section, we introduce an extension of this multiframe algorithm, which considers the non-stationarity inherent to the SR problem. Fig. 2 presents a flow chart representation of this multiframe ForWarD (MForWard) SR process.

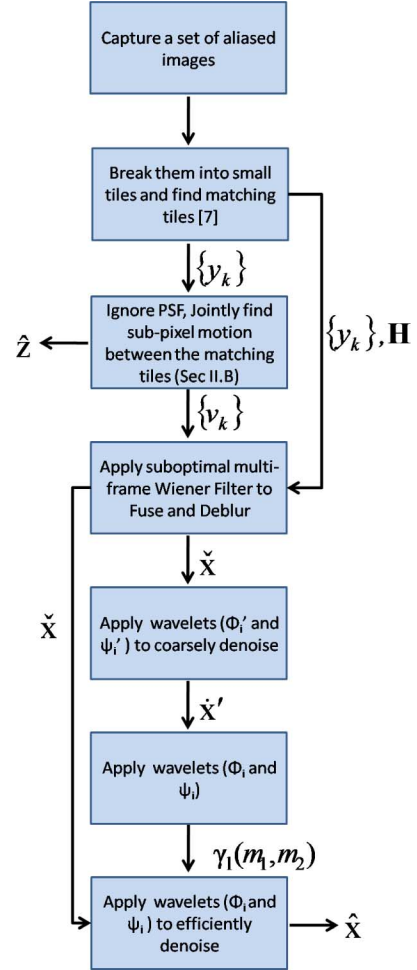


Fig. 2. Flow chart representation of the MForWard algorithm.

A. MForWard

The efficiency of the original ForWarD algorithm [5], as well as the more recent multiframe version [24], is founded on the stationarity of the noise as well as the spatial invariance of the blurring operators. The stationarity assumption breaks down in the case of multiframe SR.

In the case of SR, we must deconvolve the collection of blurry and aliased images onto a higher resolution sampling grid. The multiframe Wiener filter producing a sharp estimate of \mathbf{x} is a variant of (21) given by

$$\hat{\mathbf{x}} = \mathbf{B}^{-1}(\hat{\mathbf{y}}) \mathbf{H}^T \mathbf{g}(\hat{\mathbf{y}}) \quad (28)$$

where

$$\mathbf{B}(\hat{\mathbf{y}}) = \mathbf{H}^T \mathbf{Q}(\hat{\mathbf{y}}) \mathbf{H} + \lambda \mathbf{C}_x^{-1} \quad (29)$$

and \mathbf{Q} and \mathbf{g} are defined in (8) and (9). Here, we assume that the motion vectors $(\hat{\mathbf{y}})$ are estimated from (10). Similar to the ForWarD algorithm, we typically use values of $\lambda < 1$, which tend to sharpen the images at the expense of increased noise amplification and ringing artifacts in the resolution-enhanced image $\hat{\mathbf{x}}$. After applying the multiframe resolution enhancement

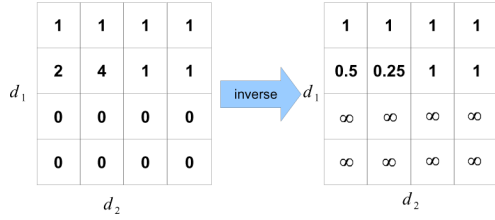


Fig. 3. Grid on the left shows the 4×4 grid of HR image sample locations and the number of LR measurements (assuming integer sampling offsets) per sample location (no blur case as in the shift-and-add image reconstruction described in [9]). In the sharpened image $\tilde{\mathbf{x}}$, the locations with fewer measurements will have higher residual noise variance. For example, the grid on the right shows the residual noise variance for an image reconstructed using an unregularized inverse of the system. The locations without any measurements have infinite noise variance.

filter of (28), the covariance matrix of the residual noise error $\tilde{\mathbf{e}}$ is

$$\mathbf{C}_{\tilde{\mathbf{e}}} = \mathbf{B}^{-1}(\hat{\mathbf{v}})\mathbf{H}^T\mathbf{Q}(\hat{\mathbf{v}})\mathbf{H}\mathbf{B}^{-1}(\hat{\mathbf{v}}). \quad (30)$$

Effective denoising and artifact removal using the wavelet shrinkage method requires accurate estimates of the residual noise power in the different wavelet filter domains. The noise covariance matrix in the wavelet coefficient domain of the wavelet function Ψ_l is given by $\Psi_l^T \mathbf{C}_{\tilde{\mathbf{e}}} \Psi_l$. In the case of the stationary model, the residual noise power σ_l^2 is constant along the diagonal of this covariance matrix. Estimating this single value is computed efficiently in (24) using FFT operations. Unlike the standard ForWarD algorithm, the multiframe estimate of the deblurred image $\tilde{\mathbf{x}}$ has a residual random error field which is *not* stationary. In the case of multiframe reconstruction, ignoring the border pixels, the residual noise field $\tilde{\mathbf{e}}$ is periodic with a period of D . For each wavelet filter, we must compute the residual noise powers $\{\sigma_l^2\}_{d_1, d_2}$ along the diagonal of $\Psi_l^T \mathbf{C}_{\tilde{\mathbf{e}}} \Psi_l$ which correspond to the D^2 HR grid locations. We use $(d_i \in \{1, 2, \dots, D\}, i \in \{1, 2\})$ to index the HR sampling locations (note Fig. 3).

The spatially-varying residual noise power depends on the collection of motion vectors $\underline{\mathbf{v}}$. That is, even though the noise fields of the captured LR images \mathbf{e}_k are stationary, different pixels in the high-resolution reconstructed image $\tilde{\mathbf{x}}$ have varying amounts of data. Because of this, the residual noise powers $\sigma_l^2(m_1, m_2)$ in the wavelet domain are not uniform and cannot be computed using (24).

To see this effect more clearly, we present a simple example. Suppose that we capture twelve LR images with a downsampling factor of $D = 4$ in both the vertical and horizontal dimensions. Furthermore, suppose that the captured images are offset by integer numbers of pixels in the high-resolution grid with the number of offsets per HR grid location shown in Fig. 3. To simplify the analysis, we assume that the signal covariance matrix is given by $\mathbf{C}_{\mathbf{x}} = (1/\lambda)\mathbf{I}$ and that the imaging system is free of blur ($\mathbf{H} = \mathbf{I}$). In this case, after applying the multiframe Wiener filter, the covariance matrix of the residual noise field $\mathbf{C}_{\tilde{\mathbf{e}}}$ is a diagonal matrix. The terms along the diagonal correspond to the residual noise variance in the reconstructed image.

The HR pixels in the sharpened image $\tilde{\mathbf{x}}$ lacking LR measurements will have much higher noise variance as indicated on the right side of Fig. 3. Indeed, in the more general case, the motion vectors do not fall perfectly onto grid locations.

Once we obtain the D^2 spatially-varying noise powers, we employ the same pattern of the traditional ForWarD algorithm using the coarse denoising by the hard thresholding of (25) to estimate the wavelet coefficient energies followed by the soft thresholding of (23). The only difference being that when we apply the hard or soft thresholding of (23) and (25), we do not use the same noise variance for every pixel. Instead, we use the spatially-varying noise powers corresponding to the HR grid locations. In this way, we incorporate the nonstationary SNR properties of the sharpened image $\tilde{\mathbf{x}}$ when performing the wavelet denoising. The next subsection explains a computationally efficient approach to this issue.

B. Efficient Fourier MForWarD

One key advantage of the original ForWarD algorithm is its efficient implementation by way of FFT operations. Using FFTs to compute (22) and (24) eliminates the need to explicitly construct the extremely large matrices to deconvolve the image and to estimate the residual noise variances. The MForWarD algorithm must also support similar computational efficiency to have any practical value. For example, computing the covariance matrix in (30) directly is computationally prohibitive due to the size of the images. We now describe an efficient implementation of the MForWarD algorithm which leverages FFT operations, analogous to the method used in the case of the stationary ForWarD approach.

We operate in the the Fourier domain as we did in Section II-C, where we originally defined many of the matrices. The only additional matrix is that of the blur operator \mathbf{H} which is spatially-invariant and, hence, is diagonal in the Fourier domain

$$\tilde{\mathbf{H}} = \text{diag}\{\tilde{H}(m_1, m_2)\}. \quad (31)$$

In the case of multiframe Wiener filtering, we again consider each spatial frequency component indexed by m'_1 and m'_2 as an independent observation model given by

$$y_k(m'_1, m'_2) = \frac{1}{D^2} \tilde{\mathbf{H}}_{m'_1, m'_2} \tilde{\mathbf{S}}_{m'_1, m'_2}(\mathbf{v}_k) \tilde{\mathbf{x}}_{m'_1, m'_2} + \tilde{\mathbf{e}}_k(m'_1, m'_2) \quad (32)$$

where

$$\begin{aligned} \tilde{\mathbf{H}}_{m'_1, m'_2} &= \text{diag}\{\tilde{H}(m'_1, m'_2), \tilde{H}(m'_1 + B, m'_2), \dots \\ &\quad \tilde{H}(m'_1 + (D-1)B, m'_2 + (D-1)B)\} \\ \tilde{\mathbf{x}}_{m'_1, m'_2} &= [\tilde{x}(m'_1, m'_2), \tilde{x}(m'_1 + B, m'_2), \dots \\ &\quad \tilde{x}(m'_1 + (D-1)B, m'_2 + (D-1)B)]^T. \end{aligned}$$

A single spatial frequency in the captured image is a function of the original signal content and summation of the $D^2 - 1$ aliased spectral components. This demonstrates that the reconstruction can be applied to each collection of aliased spectral components independently.

Estimates of the D^2 spatial frequency components in the high-resolution image from the multiple measurements of the observed spatial frequencies ω'_1, ω'_2 are obtained via

$$\tilde{\mathbf{x}}_{m'_1, m'_2} = \tilde{\mathbf{B}}_{m'_1, m'_2}^{-1}(\mathbf{y}) \tilde{\mathbf{H}}_{m'_1, m'_2}^T \tilde{\mathbf{g}}_{m'_1, m'_2}(\mathbf{y}) \quad (33)$$

where

$$\begin{aligned} \tilde{\mathbf{B}}_{m'_1, m'_2}(\mathbf{y}) &= \tilde{\mathbf{H}}_{m'_1, m'_2}^T \tilde{\mathbf{Q}}_{0,0}(\mathbf{y}) \tilde{\mathbf{H}}_{m'_1, m'_2} + \lambda \tilde{\mathbf{C}}_{\mathbf{x}, m'_1, m'_2}^{-1} \\ \tilde{\mathbf{C}}_{\mathbf{x}, m'_1, m'_2} &= \text{diag}\{\tilde{\chi}(m'_1, m'_2), \tilde{\chi}(m'_1 + B, m'_2), \dots \\ &\quad \tilde{\chi}(m'_1 + (D-1)B, m'_2 + (D-1)B)\} \end{aligned}$$

and $\tilde{\mathbf{Q}}$ is defined in (8). As in Section II-C, estimating the high-resolution image in this fashion is significantly faster than using the direct matrix form of (28). Implementing the multiframe Wiener filter in the Fourier domain requires inverting B^2 small matrices of size $D^2 \times D^2$ as opposed to inverting one very large $D^2 B^2 \times D^2 B^2$ matrix of (28). In a practical sense, any images larger than 200 by 200 pixels would require inordinate (for today's machines) amounts of processing to invert the B matrix if using the direct matrix formulation. This approach enables efficient implementation of a multiframe resolution enhancement in the Fourier domain. After estimating the sharpened image $\tilde{\mathbf{x}}$, the image is converted back into the spatial domain using an inverse FFT to obtain the multiframe filtered image in the spatial domain $\tilde{\mathbf{x}}$.

We can also apply a similar technique to accelerate the calculation of the D^2 residual noise powers in the wavelet spaces $\sigma_l^2(d_1, d_2)$, which are required for the wavelet denoising step. Because the residual noise field is not stationary, its statistics are not completely characterized by a power spectral density function. The residual noise spectral components are correlated. The $D^2 \times D^2$ covariance matrix of the residual noise field spectral components associated with the LR spatial frequency set indexed by m'_1, m'_2 is

$$\tilde{\mathbf{C}}_{\tilde{\mathbf{e}}, m'_1, m'_2} = \tilde{\mathbf{B}}_{m'_1, m'_2}^{-1}(\mathbf{y}) \tilde{\mathbf{H}}_{m'_1, m'_2}^T \tilde{\mathbf{Q}}(\hat{\mathbf{y}}) \tilde{\mathbf{H}}_{m'_1, m'_2} \tilde{\mathbf{B}}_{m'_1, m'_2}^{-1}(\mathbf{y}). \quad (34)$$

To estimate the residual noise power for a given wavelet filter Ψ_l at a particular grid location d_1, d_2 , we compute

$$\sigma_l^2(d_1, d_2) = \sum_{m'_1, m'_2} \text{Tr} \left[\tilde{\Psi}_{l, m'_1, m'_2}^T \tilde{\mathbf{S}}_{m'_1, m'_2}^T(\mathbf{d}) \tilde{\mathbf{C}}_{\tilde{\mathbf{e}}, m'_1, m'_2} \tilde{\mathbf{S}}_{m'_1, m'_2}(\mathbf{d}) \tilde{\Psi}_{l, m'_1, m'_2} \right] \quad (35)$$

where $\mathbf{d} = [d_1, d_2]^T$ and "Tr" represents the trace operator and the diagonal matrix

$$\tilde{\Psi}_{l, m'_1, m'_2} = \text{diag}\{\tilde{\Psi}(m'_1, m'_2), \tilde{\Psi}(m'_1 + B, m'_2), \dots, \tilde{\Psi}(m'_1 + (D-1)B, m'_2 + (D-1)B)\} \quad (36)$$

represents the samples of the wavelet filter spectral response. Equation (35) provides an efficient method for computing the residual noise powers in the wavelet domain for use in wavelet denoising. This step must be performed twice; once for the soft thresholding wavelet functions Ψ_1 and once for the hard thresholding wavelet set Ψ'_1 .

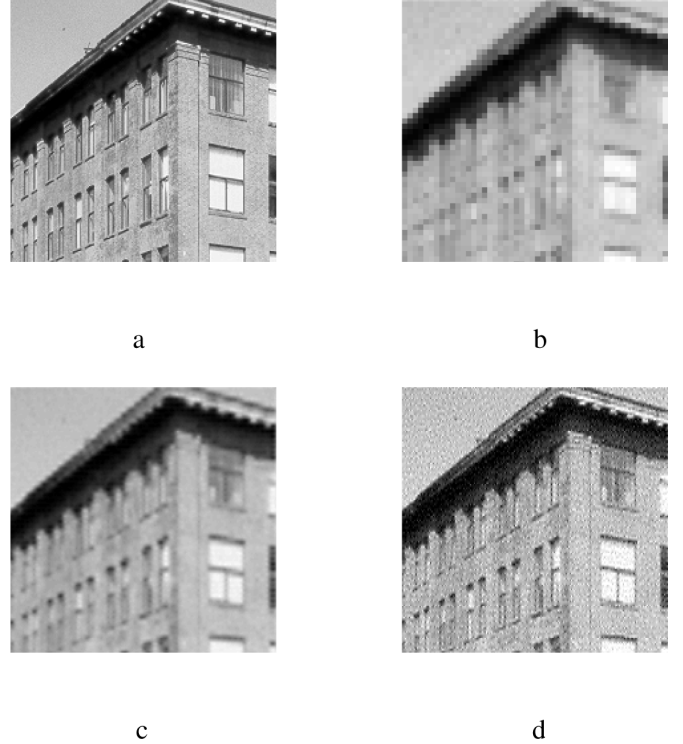


Fig. 4. Comparison of linear SR techniques for increasing the resolution of 12 simulated LR frames by a factor of $D = 4$: (a) original image; (b) captured image; (c) linear, strong prior; and (d) linear, weak prior.

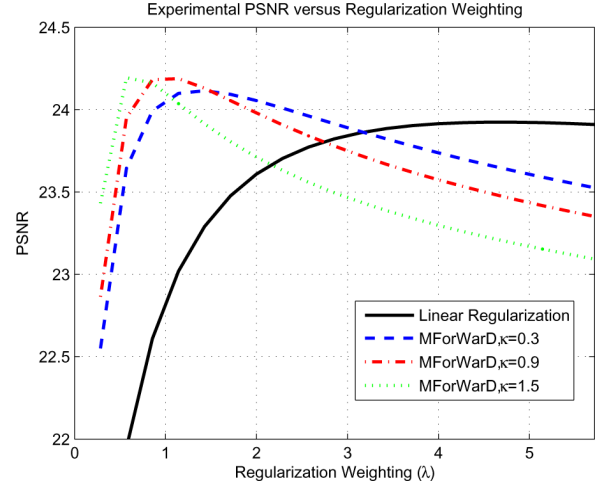


Fig. 5. Solid curve shows the PSNR performance versus regularization using standard linear regularization. The optimal weighting is around $\lambda = 4.5$ for the linear regularization. The dashed curves compare the PSNR performance versus regularization for the MForWarD algorithm using different thresholding values κ . The MForWarD algorithm shows superior peak PSNR performance over the linear regularization.

IV. EXPERIMENTAL RESULTS

In this section, we perform two sets of experiments demonstrating the capability of the MForWarD algorithm. The first section describes experiments using simulated general image data. These results demonstrate the broad applicability of the MForWarD algorithm. The second section describes some experimental results using real data captured by a digital mammogram system on a phantom breast.

A. Simulation-Based Experiments

For the first set of experiments, we construct a simple simulation example which highlights the advantages of the MForWarD algorithm for general imaging. We construct a simulated data sequence by downsampling the 400 by 400 pixel portion of a high-resolution shown in Fig. 4(a) using the subsample location shown in Fig. 3. We show only a cropped portion of the image to highlight the detail in the experiment. We simulate the optical blur using a simple heavy-tailed point spread function of the form

$$h(m_1, m_2) = \frac{1}{c} e^{-\frac{1}{\beta} \sqrt{m_1^2 + m_2^2}} \quad (37)$$

where the β term controls the rate of decay of the PSF, and c is the normalizing constant. We use a β value of 1.2 pixels. We also added noise to the LR images to create an effective SNR of about 26 dB. An example of the simulated LR image is shown in Fig. 4(b).

We reconstructed the HR image using three different approaches representing the different classes of SR algorithms. In all these simulated cases, the motion was assumed to be perfectly known. The first method is the completely linear approach embodied by the multiframe Wiener filter of (28), or the Fourier implementation of (33). This is perhaps the least computationally complex approach to multiframe SR. Using (33) produces an estimate of the HR image in a single, noniterative step and requires inverting several very small matrices. The drawback to this linear approach, however, is that the regularization of the multiframe linear filter trades off sharpness in the final image for noise reduction. For example, Fig. 4(c) shows an example of the linear reconstruction when using a power spectral density (PSD) function of the form

$$\tilde{\chi}(m_1, m_2) = \frac{\gamma^2}{|m_1|^{\alpha_1} + |m_2|^{\alpha_2}} \quad (38)$$

with $\alpha = 1.2$. The regularization weighting of (33) is $\lambda = 6$. The reconstructed image shows poor contrast to maintain minimal noise amplification. Alternatively, if we apply the linear reconstruction with a weaker prior ($\lambda = 0.02$), the reconstructed image shows improved contrast restoration at the expense of noise amplification as seen in Fig. 4(d). This is the classic tradeoff inherent to linear restoration algorithms.

The black curve of Fig. 5 shows the PSNR tradeoff as a function of regularization weighting strength. For a small λ , the PSNR is quite poor due to weak regularization of the poorly conditioned system. When λ increases beyond 4.5, the PSNR slowly degrades as the estimate becomes overly smoothed. If the signal's PSD was perfectly defined by (38), then the PSNR-maximizing weighting parameter would be near $\lambda = 1$. The dashed curves show PSNR versus regularization weighting for the MForWarD algorithm using different thresholding parameters κ . While the performance varies considerably for large values of λ , the peak PSNR near $\lambda = 1$ for the MForWarD algorithm shows reasonable stability with respect to the thresholding parameter κ . When we apply our MForWarD algorithm, we employ 2-tap Daubechies filters for the soft thresholding wavelet functions and 6-tap Daubechies filters for the coarse denoising by way of hard wavelet coefficient thresholding.

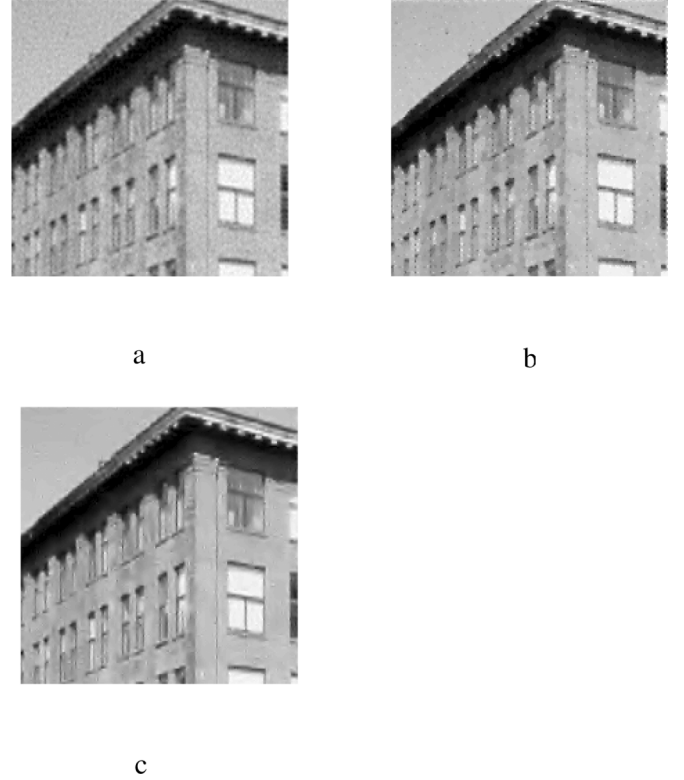


Fig. 6. Comparison of image quality for different regularization techniques using PSNR-optimal settings: (a) linear regularization (PSNR = 23.90 dB); (b) B-TV (PSNR = 23.96 dB); and (c) MForWarD (PSNR = 24.3 dB).

Fig. 6 compares the PSNR-optimal parameter settings for three different algorithms. Fig. 6(a) shows the PSNR-optimal linear regularization setting, which has a PSNR of 23.90 dB at $\lambda = 4.5$. The image shows a reasonable balance between contrast and noise gain, but still maintains a considerable amount of noise in the flat sky region. Fig. 6(b) shows the PSNR-optimal B-TV algorithm of (3) having 23.96 dB. In this experiment, $L = 2$, $\rho = 0.7$, and $\lambda = 0.0003$. The algorithm does a reasonably good job of preserving contrast while eliminating the noise in the flat regions. Fig. 6(c) shows the MForWarD algorithm at 24.3 dB using a thresholding value of $\kappa = 1.1$. The image preserves much of the contrast while eliminating the noise.

One advantage of the proposed Fourier-wavelet SR algorithm is the minimal computational overhead required for the wavelet denoising. Running on an Intel Core-2-Duo 2.2 GHz processor, the Fourier restoration requires 4.3 s of computation time. Performing the wavelet-based denoising requires an additional 2.3 s for a total of about 6.6 s. In contrast, the B-TV algorithm, representative of the large class of iterative techniques, requires only 0.5 s per iteration, but requires at least 60 steepest descent iterations (30 s) before approaching a limiting image quality.

B. Real X-Ray Data Experiments

In this section, we apply our multiframe reconstruction and restoration algorithm to real images captured on an experimental X-ray imaging system. Our experimental imaging system is based on a Mammomat NovationTOMO digital mammography prototype system (Siemens Medical Solutions,

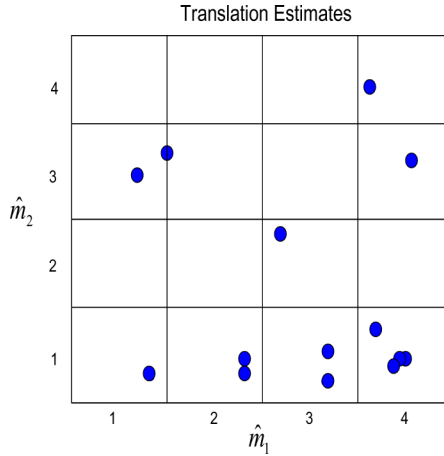


Fig. 7. Scatter plot shows the set of estimated motions $\{\hat{\mathbf{v}}_k\}$ on the HR image grid. Some high-resolution pixels have as many as four measurements whereas others have none underscoring the need for a spatially-varying denoising approach.

Erlangen, Germany),³ stationed at the Duke University Medical Center. The system uses a stationary selenium-based detector of $85\ \mu\text{m}$ pixels. Pixels with this size correspond to a Nyquist sampling rate of 5.6 line pairs per millimeter (lp/mm). We use a CIRS model 11A breast phantom (CIRS Inc., Norfolk, VA) to test our SR algorithms. We introduce shifts in the image by two methods. First, we allow the x-ray tube to rotate by ± 1 degree. Second, we manually move the breast phantom to introduce motion into the system. This manual motion is completely uncontrolled. Our dataset consists of 15 frames at the low dosage level of 11.3 mAs at 28 kVp tube voltage. As a point of reference, we also acquire a single frame at a more typical dosage of 226 mAs at 28 kVp tube voltage (Fig. 1). The breast phantom includes several testing features including a pair of resolution bar charts. We focus on the results of the test resolution chart to explore the contrast performance of the multiframe imaging system. We apply our algorithm to 100×100 pixel tiles in the captured image to estimate 400×400 pixel high-resolution images (enhancement $D = 4$).

We modeled our system PSF as a heavy-tailed exponential energy distribution with $\beta = 1.5$. To get a measure of the PSNR, we calculated the standard deviation in a textureless region of the phantom. We also measured the difference in grayscale values between for the registration bars in the resolution chart to get an approximate PSNR value of 3 dB. We fit the α terms of (38) to the periodogram obtained by averaging the spectral energy distribution over the collection of LR images. We employed 2-tap Daubechie filters for the the soft thresholding wavelet functions and 6-tap Daubechie filters for the coarse denoising by way of hard wavelet coefficient thresholding.

Fig. 7 shows a scatter plot of the set of estimated motions $\{\hat{\mathbf{v}}_k\}$ on the HR image grid. The grid reflects the number of

³Caution: Investigational Device. Limited by U.S. Federal law to investigational use. The information about this product is preliminary. The product is under development and is not commercially available in the U.S.; and its future availability cannot be ensured.

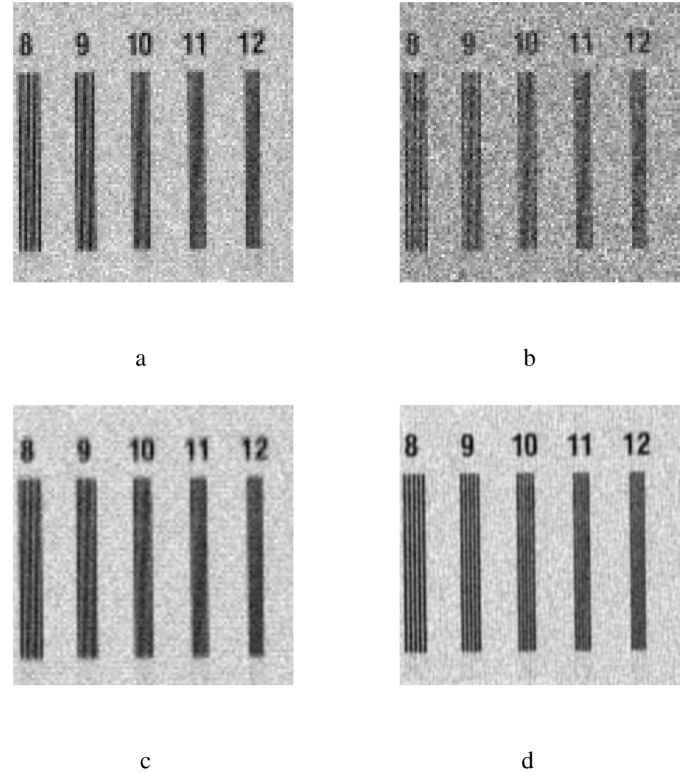


Fig. 8. (a) High dosage LR image (226 mAs); (b) low-dosage LR image (11.3 mAs); (c) motion compensated average of LR frames (no resolution enhancement); and (d) Multiframe reconstruction image $\hat{\mathbf{z}}$ of (7).

image samples per pixel in the reconstructed image. The example shows some pixel estimates combining as many as four measurements, whereas other pixels have no measurements underscoring the spatial variability of the residual noise variance.

Fig. 8 gives a visual example of the SNR for an image obtained by increasing the radiation of a single exposure versus the SNR after combining multiple low-exposure images. We focus on the portion of the resolution chart beyond the Nyquist rate for the imaging system (5.6 lp/mm). The numbers indicate the resolution in terms of line pairs per millimeter (lp/mm). The first image Fig. 8(a) shows an example of an image captured at a typically high radiation dosage of 226 mAs. The bar targets clearly show aliasing artifacts. The second image Fig. 8(b) shows the same portion of the resolution chart captured at a much lower dosage (11.3 mAs). The image demonstrates the extremely poor SNR of the captured images at such low illuminating radiation. The third image Fig. 8(c) shows the result of averaging the motion compensated LR frames without enhancing the resolution or sampling rate of the system. Interestingly, the multiframe average appears to have an approximately equivalent SNR to the single frame captured at a normal radiation dosage, although it contains some blur as a result of averaging frames with subpixel sampling offsets. The fourth image Fig. 8(d) shows the reconstructed image $\hat{\mathbf{z}}$ after registering the collection of images using the multiframe algorithm described in Section II. The reconstructed image shows a restored resolution above the Nyquist rate inherent to the detector. The image is, however, still noisy and has low contrast. The effective SNR seems comparable to that of the single image captured under high dosage Fig. 8(a).

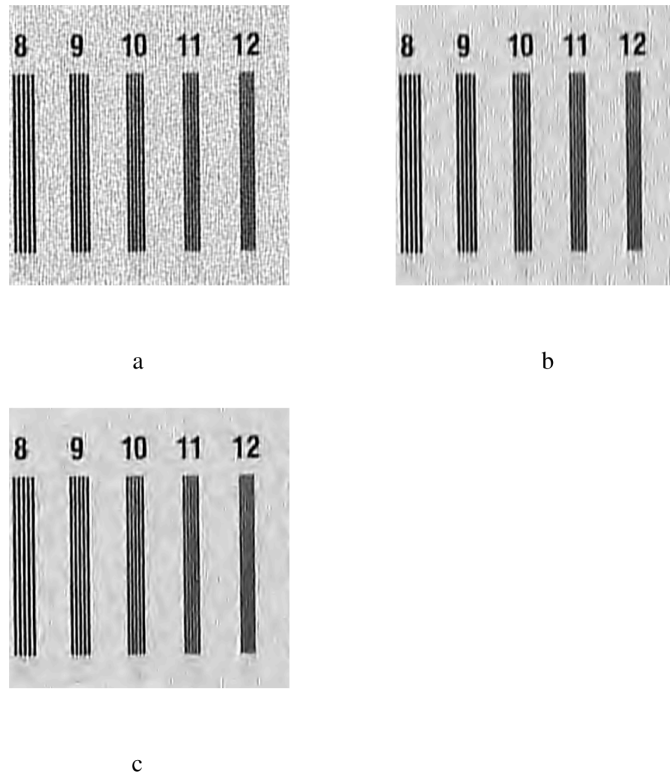


Fig. 9. Different restoration techniques applied on the low-dosage sequence illustrated in Fig. 8(b): (a) multiframe sharpened image \hat{x} ; (b) basic ForWarD \hat{x} applied to multiframe sharpened image; and (c) MForWarD \hat{x} . The MForWarD algorithm provides superior noise removal while preserving contrast.

Fig. 9 shows the resulting images after applying the sharpening and denoising steps of Section III. The left image Fig. 9(a) shows the result \hat{x} after applying the multiframe Wiener sharpening filter of (28). The image shows improved contrast with increased sharpness, but also amplified noise. The image Fig. 9(b) shows the result after the traditional ForWarD algorithm applied to the Wiener sharpened image Fig. 9(a). The hard threshold value was chosen to be $\kappa = 6$. The resulting image preserves the contrast around the bar chart signal locations while eliminating much of the noise in the signal-free portions of the image. The image still contains some residual noise. The final image Fig. 9(c) shows the result after applying the MForWarD algorithm including the nonstationary noise power computations. The hard threshold for the MForWarD algorithm was also $\kappa = 6$. Including the spatially-varying noise powers improves the noise removal over the basic ForWarD algorithm Fig. 9(b).

To get another perspective on the effects of the multiframe restoration and denoising, we plot slices through the resolution test chart region as indicated in Fig. 10 (top). Fig. 10 shows slices through several images. The top curve shows the slice through the average of the captured images containing only sub-pixel motion. The slice shows some aliasing as well as lost contrast for the bars about the Nyquist sampling rate of 5.6 lp/mm. The second graph shows a slice through the multiframe reconstructed image \hat{z} . The reconstruction eliminates the aliasing artifacts and effectively restores contrast beyond the sampling rate of the detector. The signal strength above 8 lp/mm, however, is

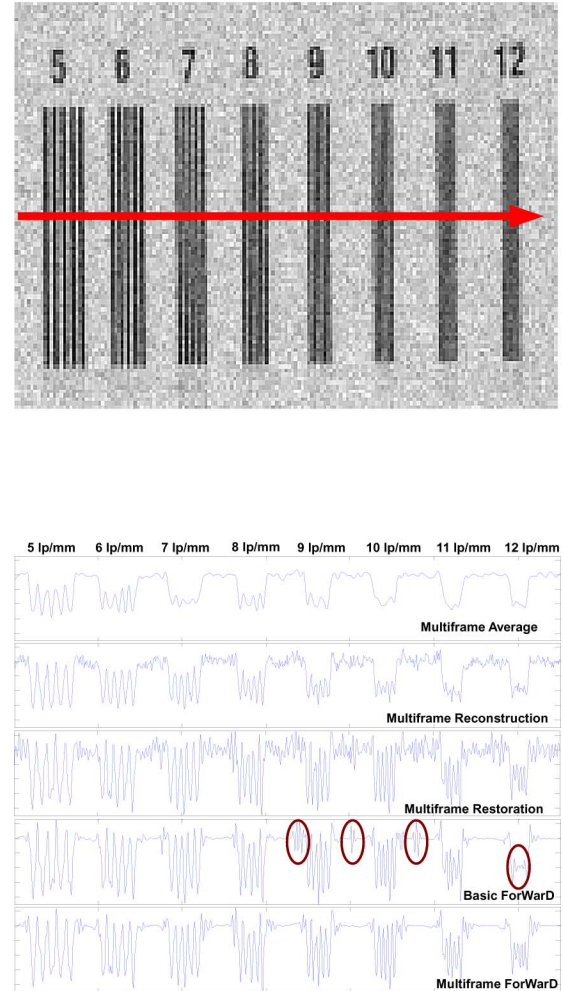


Fig. 10. Top: input low-dosage image [a zoomed in version of which was illustrated in Fig. 8(b)]. The arrow marks the sampling slice through the resolution chart. Bottom: five curves show slices through the horizontal resolution bar charts. The Nyquist rate of the system corresponds to 5.6 lp/mm. The top slice shows a slice through an interpolated average of the captured images showing aliasing artifacts and lost contrast. The second slice through \hat{z} shows enhanced resolution beyond the Nyquist rate, but poor contrast. The third slice through \hat{x} shows restored contrast but with noise amplification. The first three left ellipsoids in the fourth slice through the basic Forward reconstruction mark the amplified noise regions. The rightmost ellipsoid marks the lost resolution region. The bottom slice through the MForWarD estimate \hat{x} shows contrast preservation with significantly less noise.

very weak due to the blurring inherent to the imaging system. The third slice is from the multiframe restoration result (\hat{x}). The sharpening restores contrast out to the 12 lp/mm, more than twice the Nyquist rate, but at the expense of noise amplification. The fourth slice is from the basic Forward reconstruction, which still has some noise amplified regions. The bottom slice shows \hat{x} after multiframe wavelet denoising. We observe that the contrast is preserved while significantly eliminating the noise in between the bar chart signal regions.

The final goal of digital mammography is the detection and diagnosis of cancerous lesions in the breast tissue. The breast phantom contains small grains of calcium for predicting the diagnostic capability of the imaging system for calcifications in the breast. The calcium grains range from 400 μm down to

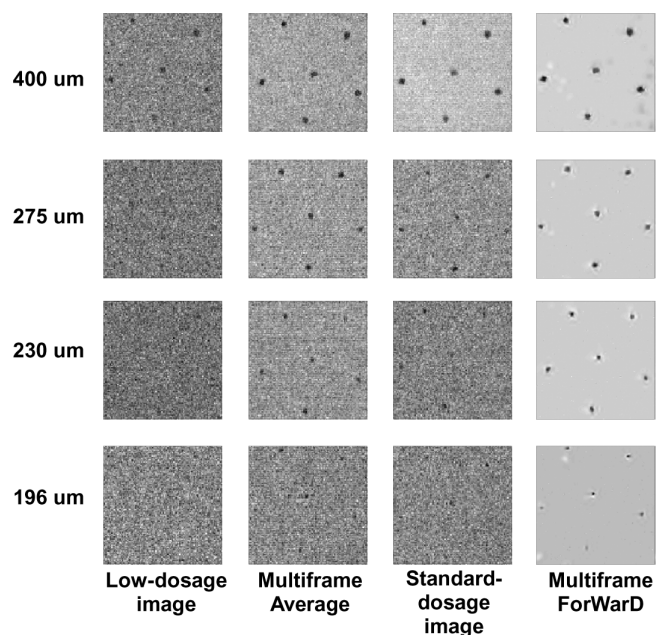


Fig. 11. Table of images shows the low-dosage image (first column), multiframe average (second column), standard dosage image (third column) and MForWarD image of calcium carbonite deposits with decreasing grain size. The synthetic calcifications are clearly visible in all of the MForWarD images.

196 μm . Fig. 11 shows the performance of the MForWarD algorithm for enhancing the contrast of these small deposits. The MForWarD algorithm enhances the contrast of even the smallest grains of calcium carbonite. In fact, the grains are visible even when the grains are nearly indistinguishable from noise in the single images captured at a standard dosage.

V. CONCLUSION

In this paper, we have proposed a novel method for restoring and denoising super-resolved low dosage X-ray images in a fast multiframe variant of the ForWarD algorithm of [5]. The proposed Fourier multiframe restoration and wavelet denoising algorithm provides high contrast super-resolved images while improving the extremely poor SNR of low-dosage images. The experimental results confirm that multiframe imaging can provide an alternative in the SNR versus resolution tradeoff for digital mammography. We note that our restoration algorithm can be easily modified to further enhance the quality of other Shift-and-Add based SR techniques [2], [29].

The design of future X-ray imaging systems would benefit from a systematic analysis of the resolution and SNR required for mammographic screening and diagnosis. Currently, there are no publicly available databases of multiframe aliased digital mammography images, only digitized analog mammography images. Upon approval of an institutional review board (IRB), we intend to apply this technique to imagery collected from patients in the clinical setting. At this point, we hope to demonstrate the ability of this technique to improve image quality on real digital mammography images.

In the future, one might explore the fundamental tradeoffs between radiation exposure, number of frames, and reconstruction performance. Furthermore, extensions to the ForWarD algorithm which include more sophisticated redundant wavelet

techniques such as curvelets [30] or ridgelets [31] might show even better performance. Recent research has shown that use of more sophisticated wavelets have been found to improve the quality in other medical imaging applications [32]. Future research might also include exploration of the effects of such multiframe imaging on the higher-level segmentation or detection tasks associated with digital mammography.

ACKNOWLEDGMENT

Our MATLAB software implementation of the noted algorithms is in part based on the ForWarD software, developed by Dr. R. D. Neelamani of the Digital Signal Processing group at Rice University (available at <http://www.dsp.rice.edu/software/ward.shtml>). The authors would like to thank Prof. P. Milanfar of the University of California Santa Cruz for collaborating with us in the original multiframe motion estimation publication.

REFERENCES

- [1] M. D. Robinson, S. Farsiu, J. Y. Lo, and C. A. Toth, "Efficient restoration and enhancement of super-resolved x-ray images," in *Proc. 15th IEEE Int. Conf. Image Processing*, Oct. 2008, pp. 629–632.
- [2] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Advances and challenges in super-resolution," *Int. J. Imag. Syst. Technol.*, vol. 14, no. 2, pp. 47–57, Oct. 2004.
- [3] S. Borman and R. L. Stevenson, "Super-resolution from image sequences—A review," in *Proc. Midwest Symp. Circuits and Systems*, Apr. 1998, vol. 5, pp. 374–378.
- [4] S. Park, M. Park, and M. G. Kang, "Super-resolution image reconstruction, a technical overview," *IEEE Signal Process. Mag.*, vol. 20, no. 3, pp. 21–36, May 2003.
- [5] R. Neelamani, H. Choi, and R. Baraniuk, "Forward: Fourier-wavelet regularized deconvolution for ill-conditioned systems," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 418–433, Feb. 2004.
- [6] E. Pisano, C. Gatsonis, E. Hendrick, M. Yaffe, J. Baum, S. Acharyya, E. Conant, L. Fajardo, L. Bassett, C. D'Orsi, R. Jong, and M. Rebner, "Diagnostic performance of digital versus film mammography for breast cancer screening—The results of the american college of radiology imaging network (ACRIN) digital mammographic imaging screening trial (DMIST)," *New Eng. J. Med.*, pp. 1773–1783, Oct. 2005.
- [7] D. Robinson, S. Farsiu, and P. Milanfar, "Optimal registration of aliased images using variable projection with applications to super-resolution," *Comput. J.*, vol. 52, no. 1, pp. 31–42, Jan. 2009.
- [8] M. Elad and Y. Hel-Or, "A fast super-resolution reconstruction algorithm for pure translational motion and common space invariant blur," *IEEE Trans. Image Process.*, vol. 10, no. 8, pp. 1186–1193, Aug. 2001.
- [9] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Fast and robust multi-frame super-resolution," *IEEE Trans. Image Process.*, vol. 13, no. 10, pp. 1327–1344, Oct. 2004.
- [10] R. Hardie, "A fast image super-resolution algorithm using an adaptive wiener filter," *IEEE Trans. Image Process.*, vol. 16, no. 12, pp. 2953–2964, Dec. 2007.
- [11] D. Robinson, S. Farsiu, J. Y. Lo, P. Milanfar, and C. A. Toth, "Efficient multiframe registration of aliased X-ray images," in *Proc. 41st Asilomar Conf. Signals, Systems, and Computers*, Nov. 2007, pp. 215–219.
- [12] R. R. Schultz and R. L. Stevenson, "Extraction of high-resolution frames from video sequences," *IEEE Trans. Image Process.*, vol. 5, no. 6, pp. 996–1011, Jun. 1996.
- [13] M. Elad and A. Feuer, "Restoration of single super-resolution image from several blurred, noisy and down-sampled measured images," *IEEE Trans. Image Process.*, vol. 6, no. 12, pp. 1646–1658, Dec. 1997.
- [14] H. Takeda, S. Farsiu, and P. Milanfar, "Deblurring using regularized locally adaptive kernel regression," *IEEE Trans. Image Process.*, vol. 17, no. 4, pp. 550–563, Apr. 2008.
- [15] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. IEEE Int. Conf. Computer Vision*, Jan. 1998, pp. 836–846.
- [16] V. Patanavijit and S. Jitapunkul, "A Lorentzian stochastic estimation for a robust iterative multiframe super-resolution reconstruction with Lorentzian-Tikhonov regularization," *EURASIP J. Adv. Signal Process.*, vol. 2007, no. 2, 2007.

- [17] N. A. El-Yamany and P. E. Papamichalis, "Robust color image super-resolution: An adaptive m-estimation framework," *J. Image Video Process.*, vol. 8, no. 2, pp. 1–12, 2008.
- [18] R. Hardie, K. Barnard, and E. Armstrong, "Joint map registration and high-resolution image estimation using a sequence of undersampled images," *IEEE Trans. Image Process.*, vol. 6, no. 12, pp. 1621–1633, Dec. 1997.
- [19] L. D. Alvarez, J. Mateos, R. Molina, and A. K. Katsaggelos, "Advances and challenges in super-resolution," *Int. J. Imag. Syst. Technol.*, vol. 14, no. 2, pp. 58–66, Oct. 2004.
- [20] D. Robinson and P. Milanfar, "Statistical performance analysis of super-resolution," *IEEE Trans. Image Process.*, vol. 15, no. 6, pp. 1413–1428, Jun. 2006.
- [21] P. Vandewalle, S. Susstrunk, and M. Vetterli, "A frequency domain approach to registration of aliased images with application to super-resolution," *EURASIP J. Appl. Signal Process.*, 2006.
- [22] N. Woods, N. Galatsanos, and A. Katsaggelos, "Stochastic methods for joint registration, restoration, and interpolation of multiple undersampled images," *IEEE Trans. Image Process.*, vol. 15, no. 1, pp. 201–213, Jan. 2006.
- [23] G. Golub and V. Pereyra, "Separable nonlinear least squares: The variable projection method and its applications," *Inst. Phys. Inv. Prob.*, vol. 19, pp. R1–R26, 2003.
- [24] R. Neelamani, M. Deffenbaugh, and R. Baraniuk, "Texas two-step: A framework for optimal multi-input single-output deconvolution," *IEEE Trans. Image Process.*, vol. 16, no. 11, pp. 2752–2765, Nov. 2007.
- [25] A. K. Jain, *Fundamentals of Digital Image Processing*, 1st ed. Upper Saddle River, New Jersey: Prentice-Hall, 1989.
- [26] A. Murli, L. D'Amore, and V. De Simone, "The wiener filter and regularization methods for image restoration problems," in *Proc. Int. Conf. Image Analysis and Processing*, Sep. 1999, pp. 394–399.
- [27] J. Fowler, "The redundant discrete wavelet transform and additive noise," *IEEE Signal Process. Lett.*, vol. 12, no. 9, pp. 629–632, Sep. 2005.
- [28] S. Mallat, *A Wavelet Tour of Signal Processing*. New York: Academic, 1998, p. 442.
- [29] H. Takeda, S. Farsiu, and P. Milanfar, "Kernel regression for image processing and reconstruction," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 349–366, Feb. 2007.
- [30] J.-L. Starck, E. Candes, and D. Donoho, "The curvelet transform for image denoising," *IEEE Trans. Image Process.*, vol. 11, no. 6, pp. 670–684, Jun. 2002.
- [31] M. Do and M. Vetterli, "The finite ridgelet transform for image representation," *IEEE Trans. Image Process.*, vol. 12, no. 1, pp. 16–28, Jan. 2003.
- [32] A. Khare and U. S. Tiwary, "A new method for deblurring and denoising of medical images using complex wavelet transform," in *Proc. IEEE Conf. Engineering in Medicine and Biology*, Sep. 2005, pp. 1897–1900.



M. Dirk Robinson received the Ph.D. degree in electrical engineering from the University of California, Santa Cruz, in 2004, where he developed super-resolution algorithms.

He is currently the Manager of the Digital Optics Research group at Ricoh Innovations, Inc., Menlo Park, CA, researching practical applications of computational imaging technology. His work ranges from optical imaging systems design to image processing algorithm research.

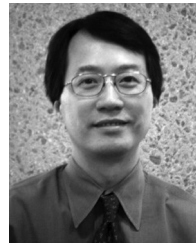


Cynthia A. Toth received the M.D. degree from the Medical College of Pennsylvania.

She completed ophthalmology residency training at Geisinger Medical Center and a Fellowship in Vitreoretinal Diseases and Surgery at University of California, Davis. She was Chief of the Retina Service at Wilford Hall USAF Medical Center. Joining Duke University Medical Center Faculty, Durham, NC, as a vitreoretinal surgeon in 1993, she is now Professor of Ophthalmology and of Biomedical Engineering. She heads the Duke Advanced Research in SDOCT

Imaging (DARSI) Laboratory. She is the author of over 100 peer-reviewed publications and multiple book chapters, and an inventor on 11 patents.

Dr. Toth is a Diplomate of the American Board of Ophthalmology, Fellow of the American Academy of Ophthalmology, and a member of the Retina Society and Club Jules Gonin.



Joseph Y. Lo received the B.S.E. and Ph.D. degrees in biomedical engineering from Duke University, Durham, NC.

He is currently an Associate Professor of radiology and biomedical engineering, faculty of the medical physics graduate program of Duke University/Duke University Medical Center. Previously, he was a postdoctoral fellow in radiology at Duke University Medical Center. His research interests span many topics in the advanced imaging and management of cancer, including breast tomosynthesis imaging, computer-

aided diagnosis and risk prediction, radiation therapy treatment planning, bioinformatics, and digital image processing.



Sina Farsiu received the B.Sc. degree in electrical engineering from Sharif University of Technology, Tehran, Iran, in 1999, the M.Sc. degree in biomedical engineering from the University of Tehran, Tehran, in 2001, and the Ph.D. degree in electrical engineering from the University of California, Santa Cruz (UCSC), in 2005.

He was a Postdoctoral Scholar at UCSC from 2006 to 2007 and a postdoctoral research associate in the Department of Ophthalmology, Duke University, Durham, NC, from 2007 to 2009. He is currently

an Assistant Professor in the Departments of Ophthalmology and Biomedical Engineering, and Fitzpatrick Institute for Photonics at Duke University, where he is the director of the Vision and Image Processing (VIP) laboratory. His technical interests include robust image enhancement and reconstruction, automatic segmentation, computer aided surgery, ocular imaging and analysis, optical coherence tomography, and digital X-ray imaging.