Wavelet-Based Deconvolution Using Optimally Regularized Inversion for Ill-Conditioned Systems

Ramesh Neelamani, Hyeokho Choi, and Richard Baraniuk

Rice University
Houston, Texas 77005, USA

ABSTRACT

We propose a hybrid approach to wavelet-based deconvolution that comprises Fourier-domain system inversion followed by wavelet-domain noise suppression. In contrast to conventional wavelet-based deconvolution approaches, the algorithm employs a *regularized inverse filter*, which allows it to operate even when the system is non-invertible. Using a mean-square-error (MSE) metric, we strike an optimal balance between Fourier-domain regularization (matched to the system) and wavelet-domain regularization (matched to the signal/image). Theoretical analysis reveals that the optimal balance is determined by the economics of the signal representation in the wavelet domain and the operator structure. The resulting algorithm is fast ($O(N \log_2^2 N)$ complexity for signals/images of N samples) and is well-suited to data with spatially-localized phenomena such as edges. In addition to enjoying asymptotically optimal rates of error decay for certain systems, the algorithm also achieves excellent performance at fixed data lengths. In simulations with real data, the algorithm outperforms the conventional time-invariant Wiener filter and other wavelet-based deconvolution algorithms in terms of both MSE performance and visual quality.

Keywords: deconvolution, restoration, wavelets, regularization

1. INTRODUCTION

Deconvolution is a recurring theme in a wide variety of signal and image processing problems, from channel equalization[1] to image restoration.[2] For example, satellite images obtained in practice are often blurred due to limitations such as aperture effects of the camera, camera motion, or atmospheric turbulence. Deconvolution becomes necessary if we wish a crisp, deblurred image for viewing or further processing.

1.1. Problem description

In its simplest form, the 1-d deconvolution problem runs as follows. The desired signal x is input to a known linear time-invariant (LTI) system $\mathcal{H}$ having impulse response h. Independent identically distributed (i.i.d.) samples of the addtive white Gaussian noise (AWGN) γ with variance σ^2 corrupt the output samples of the system $\mathcal{H}$ (see Fig. 1). The observations y at discrete points t_n are given by

$$y(t_n) := (x \circledast h)(t_n) + \gamma(n), \qquad n = 0, \ldots, N-1. \tag{1}$$

For simplicity (but without loss of generality), we assume circular convolution, denoted by $\circledast$ in (1). Given y, we seek to estimate x.

In the discrete Fourier transform (DFT) domain, we equivalently have

$$Y(f_n) = H(f_n)\, X(f_n) + \Gamma(f_n), \qquad n = 0, \ldots, N-1, \tag{2}$$

where Y, H, X and Γ are the respective length-N DFTs. The $f_n := \frac{2\pi n}{N}$ denote the normalized frequencies in the DFT domain. The problem formulation trivially extends to multi-dimensional data.

If the system frequency response $H(f_n)$ has no zeros, then an unbiased estimate of x can be obtained by inverting $\mathcal{H}$ as

$$\begin{aligned} \widehat{X}(f_n) &:= H^{-1}(f_n) Y(f_n) \\ &= X(f_n) + H^{-1}(f_n)\, \Gamma(f_n). \end{aligned} \tag{3}$$

However, if the system is ill-conditioned, i.e., if $H(f_n)$ is small at any f_n, then enormous noise amplification results during inversion, resulting in an extremely noisy, useless estimate.

This work was supported by NSF, grant MIP–9457438, DARPA/AFOSR, grant F49620–97–1–0513, Texas Instruments, and the Rice Consortium for Computational Seismic/Signal Interpretation.

Email: neelsh@rice.edu, choi@ece.rice.edu, richb@rice.edu. Internet: www.dsp.rice.edu

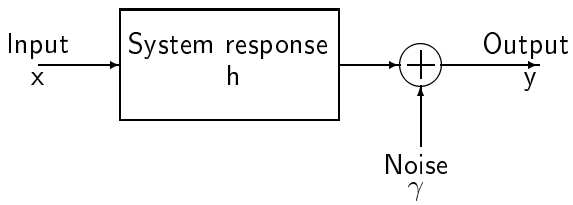


Figure 1. *Convolution model setup.*

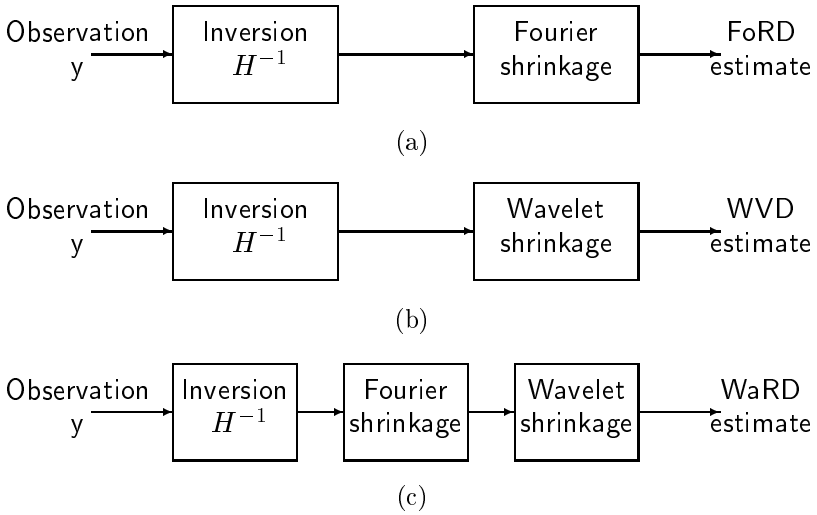


Figure 2. *Different deconvolution techniques: (a) Formally, Fourier-domain regularized deconvolution (FoRD) estimates the signal in the presence of the noise colored by H^{-1} using Fourier-domain shrinkage. Wiener deconvolution is a special case of FoRD that shrinks according to the SNR at each frequency. (b) Wavelet-vaguelette deconvolution (WVD) estimates the signal in the presence of the noise colored by H^{-1} using wavelet-domain shrinkage. (c) Our proposed wavelet regularized deconvolution (WaRD) exploits both FoRD and WVD concepts to minimize the distortion of spatially localized features in the signal.*

1.2. Fourier-domain regularized deconvolution (FoRD)

Noise amplification can be alleviated by using an approximate, regularized inverse instead of a pure inverse. Regularization aims to provide a better solution by reducing noise in exchange for some distortion in the estimate[3]; regularization becomes essential in situations involving ill-conditioned systems.

The LTI Wiener deconvolution filter is a classical example of Fourier-domain regularized deconvolution (FoRD, see Fig. 2(a)). It provides the MSE-optimal regularized LTI solution to the deconvolution problem. Formally, the estimation procedure used by the LTI Wiener deconvolution filter can be understood as comprising the inversion of the convolution operator using H^{-1} to obtain a noisy estimate $\widehat{X}(f_n)$ followed by shrinkage of each frequency component of $\widehat{X}(f_n)$ according to the signal-to-noise-Ratio (SNR) at f_n (shrink less/more when the SNR is high/low).*

The Fourier transform diagonalizes the convolution operator $\mathcal{H}$; hence the Fourier domain is ideally suited to represent the colored noise in $\widehat{X}(f_n)$ that results from its inversion. Consequently, the LTI Wiener deconvolution filter, which employs Fourier-domain noise shrinkage, can precisely identify and attenuate the noise that gets amplified during inversion of $\mathcal{H}$, thereby fully exploiting the structure of the blurring system. In fact, when the input signal can be modeled as wide-sense stationary (WSS) and Gaussian, the LTI Wiener deconvolution filter is MSE-optimal over *all* estimators.

*The inversion and shrinkage is performed jointly in practice; hence FoRD and Wiener deconvolution are applicable even when the system $\mathcal{H}$ is not invertible.

However, the Fourier domain is not well-suited for representing many common signals and images that contain spatially localized phenomena such as edges because the Fourier basis functions have support that extend over the entire spatial domain. Scalar processing of the Fourier components employed by FoRD results in lack of spatial selectivity, consequently important spatial features such as edges get distorted.

1.3. Wavelets and wavelet-vaguelette deconvolution (WVD)

Over the last decade, the wavelet transform had proven invaluable for dealing with a wide class of signals and images with spatially localized features. Wavelets provide economical representations for a large class of functions, such as those belonging to Besov spaces.[4] Wavelets capture most of the signal energy using just a few large wavelet coefficients. This property has been leveraged into powerful, spatially adaptive, signal estimation algorithms that are based on simply shrinking the wavelet coefficients of the noisy signal.

Motivated by the economy of wavelet representations, Donoho proposed the wavelet-vaguellette decomposition algorithm to solve a special class of linear inverse problems.[5] With a slight abuse of notation, we refer to the wavelet-vaguellette decomposition applied to deconvolution as wavelet-vaguelette deconvolution (WVD). In contrast to FoRD, WVD employs wavelet-domain shrinkage to estimate the signal in the presence of noise colored by H^{-1} (see Fig. 2(b)).

WVD aims to exploit the economical wavelet representation of signals to effectively identify and estimate the signal. In fact, for special classes of blurring operators such as the Radon transform, WVD exhibits asymptotically (as $N \to \infty$) near-optimal rates of error decay for a wide class of input signals.[5]

However, the wavelet transform does not diagonalize the convolution operator $\mathcal{H}$. Consequently, the noise frequency components that are amplified during inversion of $\mathcal{H}$ corrupt many wavelet coefficients. For example, for a box-car impulse response h — a commonly used model for image blurring due to camera motion[2] — the noisy estimate $\widehat{x}$ obtained after system inversion has infinite noise variance at all wavelet scales. Thus, even though wavelets provide an efficient input signal representation, signal estimation using scalar operations in the wavelet domain is futile and results in a zero signal as the estimate.

1.4. Wavelet-based regularized deconvolution (WaRD)

Motivated by the fact that the Fourier domain matches the convolution operator while the wavelet domain matches a large class of potential input signals, we propose an improved hybrid wavelet-based regularized deconvolution (WaRD) algorithm suitable for use with any ill-conditioned systems. The basic idea is simple: employ the best of *both* FoRD and WVD processing (see Fig. 2(c)). In this tandem processing, we exploit Fourier-domain regularization adapted to the convolution operator to control noise amplification. However, we use it sparingly to keep the accompanying smearing distortions to a minimum; the bulk of the noise removal and signal estimation is achieved using wavelet shrinkage.

By optimizing over an MSE metric, we find that the optimal balance between local processing with the wavelet basis and global processing with the Fourier basis is determined by both the distribution of signal energy in the wavelet domain and the convolution operator.

1.5. Related work

One extreme of our Fourier/wavelet balance is to perform no Fourier-domain regularization; this is equivalent to the WVD approach of Donoho[5] and the mirror-wavelet basis approach of Kalifa et al.[6] We show that WaRD subsumes WVD and thus WaRD possesses the same asymptotically (as $N \to \infty$) near-optimal error decay rates as WVD for special operators such as the Radon transform. However, at any fixed sample-size N, WaRD will outperform WVD. Furthermore, unlike WVD, WaRD is applicable to any convolution operator.

The mirror-wavelet basis approach of Kalifa et al.[6] adapts to the frequency response of the convolution operator $\mathcal{H}$. Though the adapted basis improves upon the WVD, it is not effective for all types of ill-conditioned systems. For example, when $\mathcal{H}$ has a box-car impulse response h, adapting to the sinc frequency response H using wavelets fails.

Nowak et al.[7] have employed an under-regularized system inverse and subsequently used wavelet-domain signal estimation. However they neither studied the implications of using the regularization nor the choice of the optimal amount of regularization.

Banham et al. apply a multiscale Kalman filter to the deconvolution problem.[8] Their approach employs an under-regularized constrained-least-squares prefilter to reduce the support of the state vectors in wavelet domain, thereby improving the computational efficiency of the multiscale restoration filter. The amount of regularization chosen for each wavelet scale is the lower bound that allows for reliable edge classification. While similar in spirit to the multiscale Kalman filter approach, in WaRD the amount of regularization is chosen to optimize the overall MSE performance of the deconvolved estimate. In addition, WaRD employs simple shrinkage on the wavelet coefficients of an over-complete wavelet basis in contrast to more complicated prediction on edge and non-edge quadtrees over an orthonormal wavelet basis.[8]

1.6. Paper organization

After discussing regularization in more depth in Section 2, we briefly review wavelet transforms and their properties in Section 3. Previous wavelet-based deconvolution approaches are discussed in Section 4. We present our improved WaRD scheme in Section 5 and elaborate on its implementation in Section 6. Illustrative examples lie in Section 7. We conclude by summarizing our work and sketching out future directions in Section 8.

2. FOURIER-DOMAIN REGULARIZED DECONVOLUTION (FoRD)

Given the general deconvolution problem from the Introduction, FoRD can be understood formally as follows (see Fig. 2(a)):

1. **Pure inversion:** The noisy and blurred observation y is treated by H^{-1} to obtain a noisy, unbiased estimate $\widehat{x}$ of the input signal x as in (3). This necessarily amplifies the noise components at frequencies where $H(f_n)$ is small.

2. **Fourier-domain signal estimation:** Each frequency component of the noisy signal $\widehat{x}$ is shrunk using frequency-dependent weights

$$R_\alpha(f_n) := \frac{|H(f_n)|^2\,|P_x(f_n)|^2}{|H(f_n)|^2\,|P_x(f_n)|^2 \;+\; \alpha\,\sigma^2}, \tag{4}$$

   where $P_x(f_n)$ is the power spectral density (PSD) of the input signal.[†] Formally, the signal estimate is given by $\widehat{X}_{\mathrm{FoRD}}(f_n) := H^{-1}(f_n) R_\alpha(f_n) Y(f_n)$.

The parameter α, called the *regularization parameter*, controls the tradeoff between the amount of noise suppression and the amount of signal distortion. Setting $\alpha = 0$ gives an unbiased but noisy estimate. LTI Wiener deconvolution corresponds to $\alpha = 1$.[9] Setting $\alpha = \infty$ completely suppresses the noise, but also completely distorts the signal ($\widehat{x}_{\mathrm{FoRD}} = 0$).

2.1. Optimality of FoRD

For Gaussian wide-sense-stationary signals, the LTI Wiener deconvolution provides the globally MSE-optimal estimate for the input, since the Fourier domain provides the ideal representation for *both*, the colored noise after inversion and the signal of interest.

2.2. Drawbacks of FoRD

The LTI Wiener filter does not provide a good estimate when the input signal comprises of spatially localized phenomena such as edges. Though the Fourier domain is still the ideal domain to represent the noise colored by H^{-1}, this domain is not well-suited to represent the signal x. The supports of the Fourier basis functions extend over the entire spatial domain. So, scalar operations on the Fourier coefficients lack spatial localization. Hence, all spatial components of the input signal are processed uniformly to result in a substantially distorted estimate.

[†] This assumes the signal x to be a stationary random process. We will rather assume x to be deterministic and substitute $P_x(f_n) = |X(f_n)|^2$ in (4).

2.3. Alternative solutions

Deconvolution techniques need to take the spatial variations of the signal into account to produce the best possible results. One such technique is the best linear estimator, the time-varying or matrix version of the Wiener inverse.[10] However, the time-varying Wiener filter is impractical, because it requires the input signal cross-correlation matrix which needs precise knowledge of the spatially varying phenomena of the signal. Further, the processing is computationally intensive ($O(N^3)$), because the estimator possesses no special structure.

3. BACKGROUND ON WAVELETS AND SIGNAL ESTIMATION

3.1. Wavelet transform

The joint time-frequency analysis of the wavelet basis efficiently captures spatially varying features in a signal. The discrete wavelet transform (DWT) represents a 1-d signal x in terms of shifted versions of a low-pass scaling function ϕ and shifted and dilated versions of a prototype bandpass wavelet function ψ.[11,12] For special choices of ϕ and ψ, the functions

$$\psi_{j,k}(t) \;:=\; 2^{j/2}\,\psi\big(2^j t - k\big)\,, \tag{5}$$

$$\phi_{j,k}(t) \;:=\; 2^{j/2}\,\phi\big(2^j t - k\big)\,, \tag{6}$$

with $j,k \in \mathbb{Z}$ form an orthonormal basis. A finite-resolution approximation x_J to x is given by[12]

$$x_J(t) \;=\; \sum_k u_{j_0,k}\,\phi_{j_0,k}(t) \;+\; \sum_{j=j_0}^{J}\sum_k w_{j,k}\,\psi_{j,k}(t), \tag{7}$$

with wavelet coefficients $u_{j,k} := \int x(t)\,\phi^*_{j,k}(t)\,dt$ and $w_{j,k} := \int x(t)\,\psi^*_{j,k}(t)\,dt$. The parameter J controls the resolution of the wavelet reconstruction x_J of x; we see that $x_\infty = x$.

For a discrete-time signal with N samples, the N wavelet coefficients $\{u_{j_0,k}, w_{j,k}\}$ of $x(t_n)$ can be easily computed using a filter bank consisting of low-pass filters, high-pass filters, and decimators. Due to the special filter bank structure the forward and inverse wavelet transform can be computed in $O(N)$ operations. For simplicity only, we will use the periodic DWT, which uses circular convolutions in its filter bank. For brevity, we will collectively refer to the set of scaling and wavelet coefficients as $\{\theta_{j,k}\} := \{u_{j_0,k}, w_{j,k}\}$. Multidimensional DWTs are easily obtained by alternately wavelet-transforming along each dimension.[11]

3.1.1. Multiresolution and time-frequency localization of wavelets

Wavelet provide a multiresolution representation of a signal, i.e., the wavelet coefficients capture the signal features at different resolution levels. In the wavelet representation notation used, j indexes the *scale* or the resolution of analysis — large j corresponds to higher resolution of analysis, while small j corresponds to the coarse scale or lowest resolution of analysis. The scale $j = J$ corresponds to the finest scale or highest resolution of analysis. In the wavelet representation, k indexes the spatial location of analysis. For a wavelet $\psi(t)$ centered at time zero and frequency f_0, the wavelet coefficient $w_{j,k}$ measures the signal content around time $2^{-j}\,k$ and around frequency $2^j\,f_0$. Thus, wavelets exhibit simultaneous spatial and frequency localization.

3.1.2. Wavelets as unconditional bases

Wavelets provide an unconditional basis for spaces of smooth functions such as Besov spaces.[4] Important spaces such as Sobolev spaces and L_p $(1 < p < \infty)$ spaces belong to the Besov scale. In essence, the unconditional basis property of wavelets means that signals belonging to such classes are characterized by only the amplitudes of the wavelet coefficients.[4,5] In contrast, the Fourier basis is not unconditional for such a wide class of signals. The wavelet coefficients of signals in a Besov space decay exponentially with scale.

The implications of this abstract notion of unconditional basis are extremely appealing. Donoho[13] shows that unconditional basis are desirable, because they typically express the signal economically by capturing most of the signal energy in just a few coefficients. Economical representations are desirable for signal estimation and compression using non-linear approximation.[14,15] In fact, unconditional basis are in a sense optimal for these tasks.[13] Furthermore, the unconditional basis property also ensures that even simple scalar operations on each wavelet coefficient are sufficient to ensure near-optimal performance.

3.2. Signal estimation by wavelet shrinkage

Wavelets provide a natural and effective solution to the problem of signal estimation in the presence of white noise.[16,17] Many real-world signals have economical wavelet-domain representations where a few large wavelet coefficients capture most of the signal energy.[11,18] However, since a wavelet transform is orthonormal and linear, the energy of white noise remains scattered over all of the wavelet coefficients. This disparity between the signal and noise representation in the wavelet domain has been exploited in a number of powerful, near-optimal, signal estimation techniques based on simply shrinking the wavelet coefficients of the noisy signal. The optimality refered to is in terms of the rate of error decay as increasingly denser observation samples are obtained.

Since wavelets provide economical representations in the presence of spatial varying features of the signal, signal estimation in the wavelet domain preserves such features more effectively than conventional LTI techniques such as estimation methods based on the windowed singular value decomposition. In fact, wavelet-domain signal estimation techniques are near-optimal for a wide class of signals including signals in the Besov space.[13]

Wavelet domain signal estimation techniques based on shrinkage can also be extended to estimate signals in the presence of colored noise. The optimality of such a wavelet-based approach now becomes dependent on the coloring of the noise in addition to the signal class.[19]

4. WAVELET-VAGUELETTE DECONVOLUTION (WVD)

Donoho studied the application of wavelets to special linear inverse problems and proposed the wavelet-vaguelette decomposition[5] that solves the problem of deconvolution in the following way (see Fig. 2(b)):

1. **Pure inversion:** Similar to Wiener deconvolution, obtain a noisy, unbiased estimate $\widehat{x}$ of the input signal as in (3).
2. **Wavelet-based signal estimation:** In contrast to Fourier-domain shrinkage employed by the Wiener deconvolution filter, WVD estimates the signal from the noisy $\widehat{x}$ obtained after pure inversion by shrinking each wavelet coefficient of the noisy signal.[‡] The variance of noise corrupting all wavelet coefficients in a particular scale is the same, but varies with different scales.[19] So scale-dependent shrinkage is employed to estimate the signal wavelet coefficients. The inverse DWT then yields the WVD deconvolution estimate from the estimated signal wavelet coefficients.

4.1. Optimality of WVD

Donoho showed that a WVD deconvolution approach is near-optimal to recover a wide class of signals (e.g., Besov spaces) when the linear operator $\mathcal{H}$ satisfies

$$h(t) \circledast y(at) = a^{\beta}\, h(at) \circledast y(t), \tag{8}$$

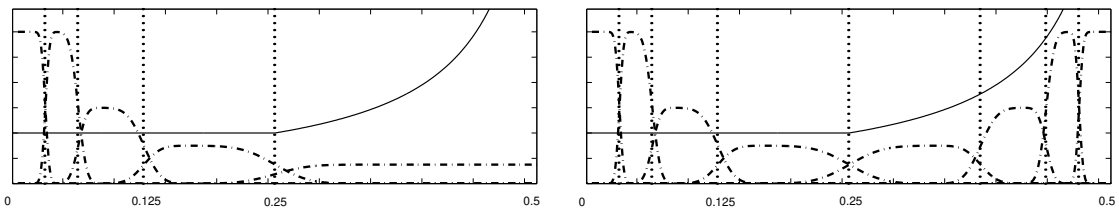
for some exponent β. Such an operator $\mathcal{H}$ is said to be *scale invariant* or *dilation homogeneous*. The Radon transform is an important example of a dilation homogeneous operator.[5] To the best of our knowledge, the optimality of wavelet-based deconvolution for general dilation-inhomogeneous operators is still unknown. Using wavelet-based techniques, Nowak et al.[7] observed impressive results for some common LTI operators as well. Kalifa et al.[6] has advocated a similar philosophy and claimed state-of-the-art performances in satellite image recovery.

We consider the example of piecewise polynomial functions to gain insight into the advantages of the WVD approach. Wavelets represent an length-N piecewise polynomial signal economically using just $O(\log_2 N)$ non-zero wavelet coefficients[§].

As a result, assuming an invertible low-pass $\mathcal{H}$, the error-per-sample decays at the rate $O(N^{\frac{-1}{2\nu+1}} \log_2 N)$ as denser sampling ($N \to \infty$) of the underlying continuous-time observations are obtained[20]; this is significantly faster than the rate $O(N^{\frac{-1}{2\nu+2}})$ achieved by Fourier-domain shrinkage. The WVD achieves faster error decay rates, because the

[‡]Note that different shrinkage techniques can be used on the noisy wavelet coefficients, but the philosophy remains the same: Use wavelet-domain estimation instead of Fourier-domain estimation

[§]Consider a wavelet system with basis functions that possess number of zero-moments that is greater than or equal to the degree of all the polynomial pieces. If the support of any wavelet basis function lies within any polynomial piece, then the corresponding wavelet coefficient is zero.



(a) (b)

Figure 3. *Deconvolution in wavelet and mirror wavelet bases. (a) The solid line is the frequency response H^{-1}. The dot-and-dashed lines show the frequency bands corresponding to wavelet basis functions at different scales. The wavelet basis functions are nearly band-limited and almost constant within dyadic frequency bands (e.g., $\frac{1}{4}$ to $\frac{1}{2}$). These dyadic frequency bands are demarcated in the figure using dotted vertical lines. If the blurring system frequency response does not have a high-order zero at normalized frequency $\frac{1}{2}$, then the variance of noise colored by H^{-1} (solid line) at different scales is primarily determined by the frequency response of the inverse within the corresponding dyadic frequency band. (b) The dot-and-dashed lines show the frequency responses of mirror-wavelet basis proposed by Kalifa et al.[6] The mirror-wavelet basis functions adopt a frequency split that aims to isolate the singularities in the inverse system frequency response and thereby reduce the noise variance in most of the mirror-wavelet basis subbands. However, if the singularity is not located at a dyadic frequency point ($\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots$) then the non-zero frequency overlap leads to infinite noise variance at all scales in any adapted wavelet basis system. The same problem occurs whenever H has a high-order zero. Thus an adapted wavelet basis system is not effective for all convolution operators.*

energy of the input signal becomes concentrated in a few large coefficients in the wavelet domain. If the noise variance in these large signal coefficients is not excessive, then identifying and retaining these signal coefficients and setting the other coefficients to zero using shrinkage gives an excellent estimate. Because the retained signal coefficients capture most of the spatial features such as edges in the signal, the final estimate retains its sharp edges as well as noise-free smooth regions.

4.2. Drawbacks of WVD

However, such a wavelet-based approach has its limitations. These can be understood by focusing on the variance of the wavelet coefficients of noise colored by H^{-1}. The noise variance σ_j^2 at scale j can be approximated by¶

$$\sigma_j^2 \approx \frac{1}{2^{(j-1)}} \sum_{n=2^{(j-1)}}^{2^{(j)}-1} \frac{1}{|H(f_n)|^2} \sigma^2, \tag{9}$$

where $H(f_n)$ are the discrete Fourier coefficients of the blurring operator $\mathcal{H}$, and σ^2 is the variance of the AWGN γ (see Fig. 3).

From (9), it is clear that if $H(f_n)$ is small at any isolated frequency f_n, then the variance of the colored noise in the entire corresponding wavelet scale explodes. Consequently, the task of extracting the signal coefficient from this scale using any form of scalar shrinking in the wavelet domain is rendered ineffective. As shown in Figure 3(a), wavelet basis functions are not exactly band-limited; the DFT coefficients of the operator outside the band $2^{(j-1)}$ to $2^j - 1$ also contribute to the noise variance but to a lesser extent. If the system frequency response H has a zero whose order is greater than the number of vanishing moments of the wavelet or if the frequency response H goes to zero at non-dyadic frequencies, then the noise variance is infinite at all the wavelet scales. In such a case, WVD provides a zero estimate.

4.3. Best-basis solution improves performance

Kalifa et al.[6] advocate adapting the wavelet basis to the frequency response of the inverse of $\mathcal{H}$ to improve on the performance of WVD. The adaptation is achieved by using a mirror-wavelet basis that possess a time-frequency tiling structure different from that possessed by conventional wavelets (see Fig. 3(b)). The tiling structure used by

¶ This approximation is valid only when the zeros in the frequency response H are of sufficiently low order.

the mirror-wavelet basis functions aims to isolate the frequency where the convolution operator H approaches zero because the variance of the noise at any scale is primarily influenced by the singularities of H^{-1} that lie in the frequency band corresponding to that wavelet scale.

However, it is not always possible to obtain a set of basis functions that can isolate all such singular frequencies in H^{-1} because the wavelet basis functions are not exactly band-limited. For example, no adapted wavelet basis scheme can provide a satisfactory solution when the blurring operator has a box-car system impulse response; the frequency response of the $\mathcal{H}$ is the sinc function. In such a case, the variance of noise colored by H^{-1} is high in all the wavelet scales, rendering signal estimation in any adapted wavelet basis domain ineffective.

5. WAVELET-BASED REGULARIZED DECONVOLUTION (WaRD)

The Fourier domian is the ideal domain to identify and attenuate noise components amplified during inversion of $\mathcal{H}$, since it diagonalizes the convolution operator. However, since the supports of Fourier basis extend over the entire spatial domain, the energy of signals containing spatially varying phenomena such as edges is spread over many Fourier coefficients. Solely employing Fourier-domain shrinkage, as in FoRD, results in significantly distorted estimates. In contrast, WVD exploits the economical signal representation in the wavelet domain. However, noise components that are severely amplified during system inversion corrupt many wavelet coefficients thereby limiting the effectiveness of wavelet-domain signal estimation.

The noise variance in the wavelet domain can be significantly reduced by attenuating the *severely amplified* Fourier components of noise. A small amount of Fourier-domain regularization achieves this without significant signal distortion. We propose a wavelet-based regularized deconvolution (WaRD) algorithm that simultaneously exploits the diagonalized Fourier-domain representation of the convolution operator and spatial adaptivity of wavelets to solve the deconvolution problem.

5.1. The WaRD algorithm

The WaRD algorithm consists of the following steps:

1. **Fourier-domain regularized inversion:** Instead of a pure inverse, employ a Fourier-domain regularized inverse (see Fig. 2(c)) to partially attenuate the amplified Fourier components of noise using weights $R_\alpha(f_n)$ (see (4)). The signal obtained after quasi-inversion is

$$\begin{aligned} \widetilde{X}_\alpha(f_n) &:= \left(\frac{1}{H(f_n)}\right) R_\alpha(f_n)\, Y(f_n) \\ &= R_\alpha(f_n)\, X(f_n) + \left(\frac{1}{H(f_n)}\right) R_\alpha(f_n)\, \Gamma(f_n), \end{aligned} \tag{10}$$

   where the parameter α, which controls the amount of Fourier-domain shrinkage employed during quasi-inversion, is typically smaller than that used to obtain a Wiener estimate ($\alpha \ll 1$). Even a small amount of regularization ensures that the severely amplified Fourier components of noise are significantly attenuated. Section 5.5 discusses the choice of α in greater depth.

2. **Wavelet-domain signal estimation:** The estimate $\widetilde{X}_\alpha$ still contains some residual noise. So, similar to step (2) in the WVD algorithm (see Section 4), shrink the wavelet coefficients of $\widetilde{X}_\alpha$ at each scale according to the noise variance at that scale to obtain the WaRD estimate (see Fig. 2(c)). Wavelet domain signal estimation remains effective since the noise corrupting the wavelet coefficients is not excessive, thanks to step 1.

5.2. Tradeoff: distortion vs. noise attenuation

Noise reduction using Fourier-domain regularization comes at the cost of signal distortion and hence needs to be controlled. In other words, this raises the question: how to pick the right value for the regularization parameter α? The tradeoff is clear: On one hand, since Fourier-domain shrinkage smears non-stationary signal features (bias) such as edges, α should be as small as possible. On the other hand, large α prevents excessive noise amplification (variance) during inversion which aids the wavelet-domain signal estimation. Thus, α controls the bias-variance tradeoff in the WaRD system.

We determine the *optimal* regularization parameter for the WaRD system by minimizing the overall MSE. This optimal regularization parameter balances the amount of noise reduction carried out in the Fourier domain and the wavelet domain.

5.3. The cost function

The overall MSE of the WaRD estimate is well-approximated by a proposed cost function that comprises of the distortion error due to the Fourier-domain regularized inversion and error incurred during wavelet-domain signal estimation. The cost function assumes that ideal thresholding T is employed during signal estimation in the wavelet domain. Ideal thresholding keeps a noisy wavelet coefficient only if the signal power in that coefficient is greater than the noise power. Otherwise, the coefficient is set to zero:

$$T\left(\widetilde{\theta}_{j,k}^{\alpha}\right) \quad = \quad \begin{cases} \widetilde{\theta}_{j,k}^{\alpha}, & \text{if } |\theta_{j,k}^{\alpha}|^2 > \sigma_j^2(\alpha) \\ 0, & \text{if } |\theta_{j,k}^{\alpha}|^2 \leq \sigma_j^2(\alpha), \end{cases} \tag{11}$$

where $\widetilde{\theta}_{j,k}^{\alpha}$ are the wavelet coefficients of the noisy signal $\widetilde{X}_{\alpha}(f_n)$ obtained after partially regularized inversion in (10), $\theta_{j,k}^{\alpha}$ are the wavelet coefficients of the distorted but noiseless input signal $R_{\alpha}(f_n)\,X(f_n)$, and $\sigma_j^2(\alpha)$ is the variance of the noise $\frac{R_{\alpha}(f_n)}{H(f_n)}\Gamma(f_n)$ at wavelet scale j. Ideal thresholding assumes that the signals under consideration are known.

The proposed cost function $\widetilde{\mathrm{MSE}}(\alpha)$ is given by

$$\widetilde{\mathrm{MSE}}(\alpha) \quad := \quad \sum_{n=0}^{N-1} [1 - R_{\alpha}(f_n)]^2\, |X(f_n)|^2 \;+\; \sum_{j,k} \min\left(|\theta_{j,k}^{\alpha}|^2,\, \sigma_j^2(\alpha)\right). \tag{12}$$

In (12), $\theta_{j,k}^{\alpha}$ denotes the wavelet coefficients of the input signal distorted by Fourier-domain regularization, $R_{\alpha}(f_n)\,X(f_n)$. The first term in $\widetilde{\mathrm{MSE}}(\alpha)$ is an estimate of the distortion in the input signal due to regularized inversion.[21] This distortion error is an increasing function of α. The second term is an estimate of the error due to ideal wavelet-domain thresholding.[22] This thresholding error is a decreasing function of α, since both the signal energy and the noise variance decrease as α increases. The optimal regularization parameter, denoted by $\alpha^{\star}$, corresponds to the minimum of $\widetilde{\mathrm{MSE}}(\alpha)$.

5.4. Accuracy of the cost function

The proposed cost function $\widetilde{\mathrm{MSE}}(\alpha)$ well-approximates the actual MSE of a WaRD system. The cost function assumes that the total error is composed of independent contributions from the distortion error incurred during the Fourier-domain regularization and the error due to subsequent ideal wavelet-domain thresholding. We analyze the following two cases and their effects on the cost function $\widetilde{\mathrm{MSE}}(\alpha)$.

Case (i) $|\theta_{j,k}^{\alpha}|^2 > \sigma_j^2(\alpha)$**:** The error contribution due to the ideal thresholding error term in $\widetilde{\mathrm{MSE}}(\alpha)$ is $\sigma_j^2(\alpha)$ (refer to (11)). So, the cost that $\widetilde{\mathrm{MSE}}(\alpha)$ associates with the estimation of $\theta_{j,k}$ is exactly equal to the actual MSE incurred during its estimation.

Case (ii) $|\theta_{j,k}^{\alpha}|^2 \leq \sigma_j^2(\alpha)$**:** The cost that $\widetilde{\mathrm{MSE}}(\alpha)$ associates with the estimation of $\theta_{j,k}$ is $|\theta_{j,k} - \theta_{j,k}^{\alpha}|^2 + |\theta_{j,k}^{\alpha}|^2$; however the actual error incurred is $|\theta_{j,k}|^2$. The associated cost is approximately equal to $|\theta_{j,k}|^2$ when the distortion $|\theta_{j,k} - \theta_{j,k}^{\alpha}|^2$ is small in comparison to $|\theta_{j,k}^{\alpha}|^2$. For small values of α, which typically are the case, the distortion is comparatively small; hence the cost function is fairly accurate; $\widetilde{\mathrm{MSE}}$ is within a factor of 2 of the actual MSE for any positive value of α.[‖]

Since $\widetilde{\mathrm{MSE}}(\alpha)$ is a good approximation of the actual error incurred (assuming ideal thresholding), it can be minimized to find the optimal regularization parameter $\alpha^{\star}$ that balances Fourier-domain regularized inversion and wavelet-domain signal estimation.

[‖] Cases that invoke such gross offsets are not generally encountered in practice.

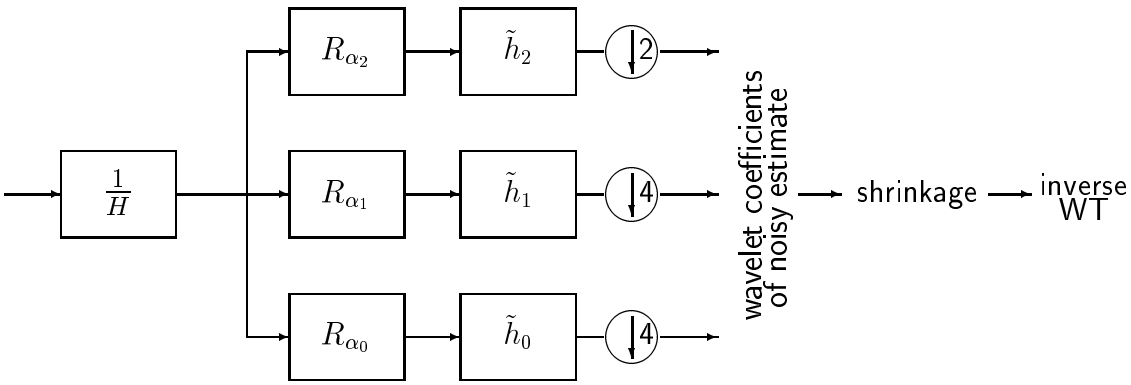


Figure 4. *Different regularization parameters for each wavelet scale: In the figure, the filters $\tilde{h}_0$, $\tilde{h}_1$ and $\tilde{h}_2$ are the equivalent, cumulative filters at each scale in a 2-level decomposition DWT filter-bank. Such an equivalent filter-bank representation can be obtained by transforming the filters in the filter bank using the noble identities*[23] *and then grouping the synthesis filters and the decimators together. Different regularization parameters, α_0, α_1, and α_2 are used at each scale.*

5.5. Optimal α for each scale

In the previous sections, we assumed a single Fourier-domain regularized inverse for all subbands. An interesting generalization is to have Fourier-domain regularized inverses with a different regularization parameters for every wavelet scale (see Fig. 4). In this case, we need to determine the optimal regularization parameter for *each* wavelet subband. Such a generalization makes the cost function separable with respect to the regularization parameter of each scale thereby simplifying the analysis. By minimizing the cost function with respect to the regularization parameter, we can show that the optimal regularization parameter $\alpha_p^\star$ for scale p satisfies[20]

$$\alpha_p^\star \approx \frac{1}{2^p} \# \left(\left| \theta_{p,k}^{\alpha_p^\star} \right|^2 > \sigma_p^2(\alpha_p^\star) \right), \tag{13}$$

where $\theta_{p,k}^{\alpha_p^\star}$ denotes the wavelet coefficients of the distorted input signal $R_{\alpha^\star}(f_n)\, X(f_n)$ at scale p, and $\sigma_p^2(\alpha_p^\star)$ denotes the variance of noise $\frac{R_{\alpha^\star}(f_n)}{H(f_n)}\Gamma(f_n)$ at scale p. In (13), $\# \left(\left| \theta_{p,k}^{\alpha_p^\star} \right| > \sigma_p^2(\alpha_p^\star) \right)$ denotes the number of wavelet coefficients $\theta_{p,k}^{\alpha_p^\star}$ that have energy greater than the noise variance $\sigma_p^2(\alpha_p^\star)$.

In words, (13) reveals that the optimal regularization parameter should be equal to the proportion of the distorted signal wavelet coefficients that lie above the noise variance. Expression (13) is a recursive expression, since both sides depend on $\alpha_p^\star$.

Consider a signal such that the signal energy in scale p is concentrated in a small proportion of the wavelet coefficients at that scale, i.e. few wavelet coefficients are large while most of the other wavelet coefficients are nearly zero. This economical distribution does not change significantly when the signal is distorted by using a small amount of Fourier-domain regularization, because the derivative of the total distortion introduced at zero regularization ($\alpha = 0$) is zero.[20] On the other hand, the rate of decrease of the noise variance with α_p in the wavelet scale is maximum at $\alpha_p = 0$. So, $\sigma_p^2(\alpha_p)$ decreases rapidly at small values of α_p. Condition (13) suggests that the noise variance be reduced using Fourier-domain regularization so that a sufficient number of wavelet coefficients of the signal are larger than the noise variance.

The optimal regularization parameter $\alpha_p^\star$ is never zero. If $\alpha_p = 0$ satisfied (13), it would imply that the noise variance is greater than the energy of each wavelet coefficient. Hence all wavelet coefficients at scale p would be shrunk to zero during wavelet-domain estimation. For most real world signals, a significant proportion of the wavelet coefficients are ≈ 0; so $\alpha_p = 1$ also will not satisfy (13). Hence $\alpha_p = 1$ (Wiener filter) also is not an optimal choice.

Several values of α_p could potentially satisfy the optimality criterion. The choice of the optimal regularization parameter in such a case is not clear. However, choosing any α_p sufficiently greater than zero (e.g., $\alpha_p \approx 0.2$) was

found to provide estimates comparable to those obtained by choosing the optimal α when the true signal spectrum $|X(f_n)|^2$ was used during Fourier-domain regularized inversion.

Thus, the expression for the optimal regularization parameter quantifies the notion of balancing Fourier-domain inversion and wavelet-domain estimation within a WaRD.

5.6. Optimality of WaRD

WVD[5] and the mirror-wavelet basis technique[6] are special cases of WaRD** with $\alpha = 0$. By construction, WaRD includes the value $\alpha = 0$ in the search-space for the optimal $\alpha^\star$. Hence WaRD enjoys all the desirable properties of Donoho's WVD such as optimal rate of error decay for dilation-homogeneous operators.

The optimal $\alpha^\star$ is never zero at a finite resolution N (though it may approach zero with increasing N). Hence, WaRD will outperform wavelet-based deconvolution methods described in[5, 6, 24] in terms of MSE at a given resolution.

Techniques based on the WVD[5, 6] (as described in Section 4) are in general not applicable when $\mathcal{H}$ is not invertible. However, thanks to the FoRD aspect of WaRD, WaRD gives excellent estimates even when $\mathcal{H}$ is non-invertible.

6. WaRD IMPLEMENTATION

The WaRD algorithm as outlined in Section 5 assumes knowledge of the variance σ^2 of the additive noise γ and the Fourier spectrum $|X(f)|^2$ of the input signal. However, these are typically unknown in practice and hence need to be estimated. The variance of the additive noise can be reliably estimated from the observation y using a median estimator in the finest wavelet scale.[16] To estimate $|X(f)|^2$, we employ the iterative Wiener technique.[25] However, since this estimation is not robust at frequencies where $H(f_n) \approx 0$, we terminate the algorithm after 10 iterations and add a small positive constant to the estimate thereby boosting the estimated $|X(f)|^2$ at high frequencies.

Condition (13) provides the necessary condition for the regularization parameter to optimally balance Fourier-domain regularization and wavelet-domain estimation at each scale. However, (13) cannot be used in practice because it assumes knowledge of the distorted wavelet coefficients of the *unknown original signal*. Further, the derivation for the optimal regularization parameter assumes ideal thresholding, which cannot be implemented in practice. In practice, the optimal regularization parameter is also influenced by the wavelet-domain estimation algorithm (hard thresholding, soft thresholding, wavelet-domain Wiener filtering, etc.) used in the WaRD system. So, we estimate the regularization parameter that is common for all wavelet scales empirically from a plot of the norm of the WaRD estimate versus the amount of regularization.[20]

A variety of wavelet-domain estimation techniques can be used in practice. The choice of a good wavelet-domain estimation scheme influences the final performance significantly. Recently, the wavelet-domain Wiener shrink estimation algorithm was proposed by Ghael et al.[26] and analyzed further by Choi et al.[27] It was observed that this technique outperforms in the MSE sense the conventional wavelet-domain estimation schemes that employ hard and soft thresholding. Estimation using a wavelet-domain Wiener shrinkage consists of the following steps: First obtain a rough estimate of the input signal using conventional wavelet-domain thresholding techniques. Then, use this estimate to obtain a final refined estimate by employing Wiener estimation on each *wavelet* coefficient.

Wavelet domain estimation schemes that use the DWT are not shift-invariant, shifts of y will result in different estimates. The redundant, shift-invariant DWT yields significantly improved estimation[28] by simply averaging over all possible shifts of the observation y at no significant increase in the overall computational cost. We employ a redundant, shift-invariant DWT with a wavelet-domain Wiener shrinkage to estimate the input signal. The computational complexity of calculating the redundant DWT is $O(N \log N)$, where N is the number of samples. For a full decomposition, there are $\log(N)$ wavelet scales. Calculation of the variances at each scale requires computing the convolution of the regularized inverse with the cumulative wavelet filter at that scale. Thus, the overall computational complexity of the algorithm, given the regularization parameter and $|X(f_n)|^2$, is $O(N \log^2 N)$.

**For comparisons with the mirror-wavelet basis approach,[6] the use of a similar adapted wavelet basis would be required.

7. RESULTS

We illustrate the performance of the WaRD algorithm using a 2-d simulation as described by Banham et al.[8] The input x is the 256×256 Cameraman image and the discrete-time system response h is a 2-d, 9×9-point smoother. Such a response is commonly used as a model for blurring due to a square scanning aperture such as in a CCD camera.[2] The blurred SNR (BSNR) is defined as $10 \log_{10} \left(\|x \circledast h\|_2^2 / N\sigma^2 \right)$. The additive noise variance was set such that the BSNR is 40 dB. Figure 5 illustrates the desired x, the observed y, the Wiener filter estimate $\widehat{x}_1$, and the WaRD estimate for $\alpha^\star = 0.1$ (determined empirically[20]). The WVD[5] and mirror wavelet basis[6] methods are not applicable in this situation, due to the many zeros in frequency response of the blurring operator $\mathcal{H}$. The WaRD estimate is clearly better that the Wiener estimate in overall visual quality and MSE. The improvement in the SNR (ISNR) is defined as $10 \log_{10} \left(\|x - y\|_2^2 / \|x - \widehat{x}\|_2^2 \right)$ where $\widehat{x}$ is the estimate. For the same experiment, Banham et al. report an ISNR of 6.68 dB using their the multiscale Kalman filter.[8] In contrast, the proposed WaRD technique provides an ISNR of 10.6 dB. The visual quality of the WaRD estimate as compared to their techique is also better (see Figure 7(d) in Banham et al.[8]).

The difference in the quality of the estimates obtained using the Wiener filter and WaRD can be grasped by viewing the cross-sections through the images. Figure 6 shows the cross-sections through row 160 of all images. The Wiener estimate cross-section shown in Figure 6(c) illustrates the failure of the Wiener technique to adapt to the smooth regions and the edges in the image simultaneously. This lack of spatial-localization reflects as ripples in the Wiener estimate. In contrast, Figure 6(d) clearly illustrates the spatial-adaptivity of WaRD. We observe the smooth regions and the edges are simultaneously preserved.

8. CONCLUSIONS

In this paper, we have proposed an efficient multi-scale deconvolution algorithm WaRD that optimally combines Fourier-domain regularized inversion and wavelet-domain signal estimation. The WaRD can be potentially employed in a wide variety of applications such as satellite imagery, seismic deconvolution, and channel equalization to obtain enhanced deconvolution estimates.

For spatially varying signals, the WaRD outperforms the LTI Wiener filter and WVD in terms of both visual quality and MSE performance. Since WaRD subsumes WVD, WaRD also enjoys asymptotically near-optimal rates of error decay with increasing samples for convolution operators such as Radon Transform. In addition, WaRD also improves on the performance of the WVD at any fixed resolution. Furthermore, WaRD continues to provide a good estimate of the original signal even in the presence of any non-invertible system. The computational complexity of the WaRD algorithm is just $O(N \log_2^2 N)$, where N is the number of samples.

Theoretical analysis of the ideal WaRD algorithm reveals that the optimal regularization parameter at each wavelet scale is determined by the proportion of distorted input signal wavelet coefficients that are greater than the variance of noise colored by regularized inversion. From (13), it follows that for finite data samples, inversion without Fourier-domain regularization in a wavelet-based deconvolution system is never optimal. Further, using a regularization parameter $\alpha = 1$, which corresponds to employing a Wiener deconvolution filter for inversion, is also sub-optimal. In essence, the optimal regularization parameter is simultaneously determined by the frugality of the wavelet representation of the input signal *and* the Fourier-domain structure of the convolution operator.

Since the theoretical analysis assumes knowledge of the input signal, expression (13) cannot be used in practice to determine the regularization parameter. However, fortunately, the final performance is observed to be quite insensitive to changes in the value of the regularization parameter around the optimal value. As a guide, in simulations spanning many real-world images and convolution systems, the optimal regularization parameter $\alpha^\star$ almost always lay in the range $[0.2, 0.3]$ when the true spectrum $|X(f_n)|^2$ was available. However, if the true spectrum is not available, then $\alpha^\star$ is dependent on the quality of the spectral estimate and needs to be determined empirically.

There are several avenues for future WaRD related research. We have focused on scalar processing during wavelet-domain estimation. However, there exist dependencies between the wavelet coefficients that can be exploited. We are currently working towards combining WaRD concepts with the hidden Markov tree model-based wavelet estimation.[29] We believe that exploiting such inter-dependencies in the wavelet domain will help preserve edges and other spatially localized phenomena better consequently leading to better deconvolution estimates.

One interesting twist to the approach to wavelet-based deconvolution is to first exploit the wavelet domain to estimate $x \circledast h$ from the noisy observation y and then invert the convolution operator. This technique, called the

(a) (b)

(c) (d)

Figure 5. *(a) Cameraman image x (256×256 samples). (b) Observed image y (BSNR = 40 dB). Smoothed by a 9×9-point smoother + noise. (c) Wiener filter estimate using the estimated $|X(f_n)|^2$ (SNR = 20.6 dB, ISNR = 8.8 dB). The ripples in the image result because the Fourier basis used by the Wiener filter (FoRD with $\alpha = 1.0$) have support over the entire spatial domain. (d) WaRD with $\alpha = 0.1$ (SNR = 22.4 dB, ISNR = 10.6 dB). In contrast to the Wiener estimate, the smooth regions and most edges are well preserved in the WaRD estimate, thanks to the spatially-localized wavelet basis functions. However, some faint features such as the grass are lost during wavelet-domain estimation.*

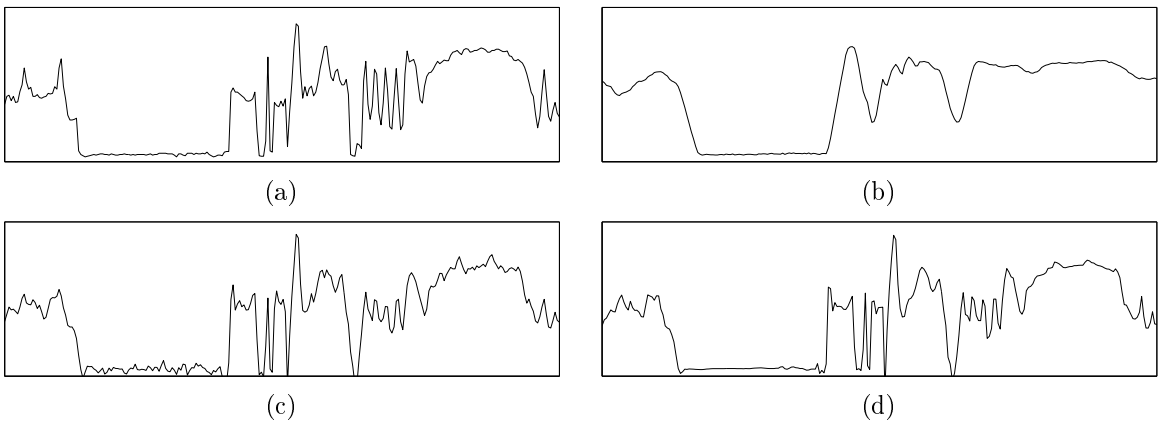


Figure 6. *(a) Cross-section of original image x (row 160 from Figure 5(a)) contains both smooth regions and discontinuities. (b) Cross-section of the blurred and noisy observed image y. (c) Cross-section of estimate obtained using the spatially invariant Wiener filter (FoRD with $\alpha = 1.0$). The Wiener deconvolution estimate exhibits ringing artifacts. (d) Cross-section of the hybrid WaRD estimate. Controlled Fourier-domain regularization ensures that residual noise can be effectively tackled by subsequent spatially-adaptive wavelet-domain estimation. WaRD preserves smooth regions and edges simultaneously even when the blurring system $\mathcal{H}$ is non-invertible.*

vaguellete-wavelet decomposition (VWD) has been studied by Silverman and Abramovich.[30] The salient point of such a technique is that the wavelet-domain estimation now deals with white noise instead of more complicated colored noise. However, in such a case the reconstruction basis is no longer a signal-adapted wavelet-basis, but rather a hybrid basis that is not spatially localized. Further, this technique is again not applicable when $\mathcal{H}$ is not invertible. Construction of a universally applicable, hybrid deconvolution scheme that lies between WVD and VWD seems both promising and challenging.

ACKNOWLEDGMENTS

We would like to thank R. Nowak for many productive discussions. A big thanks is also due to J. Romberg for helping with the WaRD implementation.

REFERENCES

1. J. G. Proakis, *Digital Communications*, McGraw-Hill, 1995.
2. A. K. Jain, *Fundamentals of Digital Image Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1989.
3. A. N. Tikhonov and V. Y. Arsenin, *Solutions of ill-posed problems*, V. H. Winston & Sons, Washington, D.C., 1977.
4. R. A. DeVore, B. Jawerth, and B. J. Lucier, "Image compression through wavelet transform coding," *IEEE Trans. Inform. Theory* **38**, pp. 719–746, Mar. 1992.
5. D. L. Donoho, "Nonlinear solution of linear inverse problems by Wavelet-Vaguellete Decomposition," *App. Comp. Harmonic Anal.* **2**, pp. 101–126, 1995.
6. J. Kalifa, S. Mallat, and B. Rougé, "Image deconvolution in mirror wavelet bases," in *Proc. IEEE Int. Conf. Image Processing — ICIP '98*, pp. 565–569, (Chicago), Oct. 1998.
7. R. D. Nowak and M. J. Thul, "Wavelet-Vaguelette restoration in photon-limited imaging," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing — ICASSP '98*, pp. 2869–2872, (Seattle, WA), 1998.
8. M. R. Banham and A. K. Katsaggelos, "Spatially adaptive wavelet-based multiscale image restoration," *IEEE Trans. Image Processing* **5**, pp. 619–634, April 1996.
9. K. R. Castleman, *Digital Image Processing*, Prentice Hall, New Jersey, 1996.
10. S. J. Orfanidis, *Optimum Signal Processing, An Introduction*, Macmillan Publishing Company, New York, NY, 1984.
11. S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, 1998.
12. C. S. Burrus, R. A. Gopinath, and H. Guo, *Introduction to Wavelets and Wavelet Transforms: A Primer*, Prentice-Hall, 1998.

13. D. L. Donoho, "Unconditional bases are optimal bases for data compression and for statistical estimation," *App. Comp. Harmonic Anal.* **1**, pp. 100–115, Dec 1993.
14. R. A. DeVore, B. Jawerth, and B. J. Lucier, "Nonlinear wavelet image processing: Variational problems, compression, and noise removal through wavelet shrinkage," *IEEE Trans. Inform. Theory* **38**, pp. 719–746, 1992.
15. D. L. Donoho, M. Vetterli, R. A. DeVore, and I. Daubechies, "Data compression and harmonic analysis," *IEEE Trans. Inform. Theory* **44**, pp. 2435–2476, Oct. 1998.
16. D. L. Donoho and I. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," *J. Amer. Stat. Assoc.* **90**, pp. 1200–1224, Dec. 1995.
17. D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation via wavelet shrinkage," *Biometrika* **81**, pp. 425–455, 1994.
18. A. Cohen and J. P. D'Ales, "Nonlinear approximation of random functions," *SIAM J. App. Math.* **57**, pp. 518–540, Apr. 1997.
19. I. M. Johnstone and B. W. Silverman, "Wavelet threshold estimators for data with correlated noise," *J. Royal Stat. Soc. B* (59), pp. 319–351, 1997.
20. R. Neelamani, "Wavelet-based deconvolution for ill-conditioned systems," M.S. thesis, Dept. of ECE, Rice Univ., in *www-dsp.rice.edu/publications/*, 1999.
21. N. P. Galatsanos and A. K. Katsaggelos, "Methods for choosing the regularization parameter and estimating the noise variance in image processing and their relation," *IEEE Trans. Image Processing* **1**, pp. 322–336, Jul. 1992.
22. D. L. Donoho, "Nonlinear wavelet methods for recovery of signals, densities, and spectra from indirect and noisy data," in *Different Perspectives on Wavelets*, vol. 47 of *Proc. Symp. Appl. Math.*, pp. 173–205, American Mathematical Society, 1993.
23. P. P. Vaidyanathan, *Multirate Systems and Filter Banks*, Prentice Hall, Englewood Cliffs, NJ, 1992.
24. R. D. Nowak, "A fast wavelet-vaguelette algorithm for discrete LSI problems," tech. rep., Michigan State University, Aug. 1997.
25. A. D. Hillery and R. T. Chin, "Iterative Wiener filters for image restoration," *IEEE Trans. Signal Processing* **39**, pp. 1892–1899, Aug. 1991.
26. S. Ghael, A. M. Sayeed, and R. G. Baraniuk, "Improved wavelet denoising via empirical Wiener filtering," in *Proc. SPIE Int. Soc. Opt. Eng.*, vol. 3169, pp. 389–399, 1997.
27. H. Choi and R. G. Baraniuk, "Analysis of wavelet domain Wiener filters," in *IEEE Int. Symp. Time-Frequency and Time-Scale Analysis*, (Pittsburgh), Oct. 1998.
28. R. Coifman and D. Donoho, "Translation invariant denoising," in *Wavelets and Statistics*, A. Antoniades, ed., Springer, 1995.
29. M. S. Crouse, R. D. Nowak, and R. G. Baraniuk, "Wavelet-based statistical signal processing using hidden Markov models," *IEEE Trans. Signal Processing* **46**, Apr. 1998. (Special Issue on Wavelets and Filter Banks).
30. F. Abramovich and B. W. Silverman, "Wavelet decomposition approaches to statistical inverse problems," *Biometrika* **85**, pp. 115–129, Oct. 1998.