WAVELET-BASED DECONVOLUTION FOR ILL-CONDITIONED SYSTEMS

Ramesh Neelamani, Hyeokho Choi and Richard Baraniuk

Department of Electrical and Computer Engineering, Rice University, Houston, TX 77251–1892, USA

ABSTRACT

In this paper, we propose a new approach to wavelet-based deconvolution. Roughly speaking, the algorithm comprises Fourier-domain system inversion followed by wavelet-domain noise suppression. Our approach subsumes a number of other wavelet-based deconvolution methods. In contrast to other wavelet-based approaches, however, we employ a regularized inverse filter, which allows the algorithm to operate even when the inverse system is ill-conditioned or non-invertible. Using a mean-square-error metric, we strike an optimal balance between Fourier-domain and wavelet-domain regularization. The result is a fast deconvolution algorithm ideally suited to signals and images with edges and other singularities. In simulations with real data, the algorithm outperforms the LTI Wiener filter and other wavelet-based deconvolution algorithms in terms of both visual quality and MSE performance.

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]. Often, the distortion introduced by a measurement device can be modeled as a convolution of the desired data with the impulse response of the device. Deconvolution, then, corresponds to inverting the effects of the distortions. Unfortunately, the measured signal is usually also corrupted by noise, which complicates the process of deconvolution.

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 having impulse response h. White Gaussian noise n of variance σ^2 corrupts the output of the system. We measure the result $y(t) := (h * x)(t) + n(t)$. In the Fourier domain, we have $Y(f) = H(f)\,X(f) + N(f)$. Given y, we seek to estimate x.

If the system frequency response $H(f)$ has no zeros, then we can obtain an unbiased estimate of x as $\widehat{X}(f) := H^{-1}(f)\,Y(f) = X(f) + H^{-1}(f)\,N(f)$. However, if $H(f)$ becomes small at any frequency, then enormous noise amplification results, yielding an infinite-variance, useless estimate.

In situations involving such ill-posed systems, some amount of *regularization* becomes essential. Regularization reduces the variance of the signal estimate (noise reduction) in exchange for an increase in bias (signal distortion). When x is wide-sense stationary (WSS), the LTI *Wiener filter* provides the optimal regularization in the minimum mean-squared-error (MSE) sense [2].

Unfortunately, the signals and images appearing in many important applications contain information-bearing edges and ridges. The LTI Wiener filter is inappropriate for such non-stationary signals, for since it reduces noise by smoothing all signal components uniformly; it can smear local features such as edges. The root of this problem lies in the fact that while the underlying Fourier basis of the LTI Wiener filter matches an LTI system h, it does not match a non-stationary signal x.

Wavelets, on the other hand, provide a basis matched to a large class of non-stationary signals [3, 4]. Unlike the strict frequency localization of the Fourier basis, a wavelet basis localizes in both time *and* frequency, and so can effectively track signal non-stationarities. This matching property has been leveraged into powerful algorithms for noise reduction that simply threshold the wavelet representation of the noisy signal [5, 6]. Wavelet denoising is a spatially adaptive processing (it smoothes more in smooth regions of the signal) ideally suited to signals with edges and other singularities.

The fact that LTI systems are matched by one basis (Fourier) but non-stationary signals by another (wavelet) inspires a hybrid approach to deconvolution: (i) invert the convolution operation in the Fourier domain and then (ii) regularize (denoise) in the wavelet domain. Such an approach has been followed by Donoho [7], Nowak [8], and Mallat [3, pp. 456-461] with considerable success.

However, current wavelet-based deconvolution schemes cannot deal with ill-conditioned systems h, since they entrust all of the regularization to their wavelet denoising post-processing step. Systems with zeroes in the frequency response will present noise of infinite variance to the denoising step, destroying the signal estimate.

In this paper, we propose an improved hybrid Fourier/wavelet deconvolution algorithm suitable for use with ill-conditioned systems. The basic idea is simple: employ *both* Fourier-domain (Wiener-like) and wavelet-domain regularization. With this tandem processing, we can keep the Fourier-domain regularization (and its corresponding smearing distortions) to the minimum required to make the system transfer function well-posed; the bulk of the noise removal comes in the wavelet denoising stage. Using the MSE metric, we will strike an optimal balance between global and local processing. Interestingly, one extreme of the balance is to perform no Fourier-domain regularization, and this coincides to the approaches of [3, 7, 8]. Extension to multi-dimensional data is trivial.

After discussing regularization in more depth in Section 2 and previous Fourier/wavelet deconvolution approaches in Section 3, we present our improved scheme in Section 4. Illustrative examples lie in Section 5. We close with conclusions in Section 6.

2. REGULARIZED INVERSE FILTERS

Consider a zero-mean, WSS signal x with power spectral density (PSD) $P_x(f)$. Given the general deconvolution problem from the

This work was supported by the National Science Foundation, grant MIP–9457438, DARPA/AFOSR, grant F49620–97–1–0513, Texas Instruments, and the Rice Consortium on Computational Seismic Interpretation.
Email: neelsh@rice.edu, choi@ece.rice.edu, richb@rice.edu
Web: www.dsp.rice.edu

Introduction, a general form for a Fourier-domain-regularized signal estimate is given by [9]

$$\widehat{X}_\alpha(f) \;:=\; G_\alpha(f)\, Y(f) \tag{1}$$

with

$$G_\alpha(f) \;:=\; \left(\frac{1}{H(f)}\right)\left(\frac{|H(f)|^2\, P_x(f)}{|H(f)|^2\, P_x(f) \,+\, \alpha\sigma^2}\right). \tag{2}$$

The regularization parameter α controls the tradeoff between the amount of noise suppression and the amount of signal distortion. Setting $\alpha = 0$ gives an unbiased but noisy estimate. Setting $\alpha = \infty$ completely suppresses the noise, but also totally distorts the signal ($\widehat{x}_\infty = 0$). For $\alpha = 1$, (2) corresponds to the LTI Wiener filter, which is optimal in the MSE sense for a Gaussian input signal x.

Since the Fourier basis functions underlying any LTI filter have spatial support over the entire signal, G_α will tend to smear non-stationarities in the desired signal, such as edges and ridges. While for non-stationary signals we could just solve the more general MSE signal estimation problem (time-varying Wiener filtering), such an approach would have no special structure and furthermore would require precise knowledge of the non-stationarities in x.

3. WAVELETS AND DECONVOLUTION

The joint time-frequency analysis of the wavelet basis efficiently captures non-stationary signal features. The discrete wavelet transform (DWT) represents a 1-d signal z in terms of shifted versions of a low-pass scaling function ϕ and shifted and dilated versions of a prototype bandpass wavelet function ψ [3, 4]. For special choices of ϕ and ψ, the functions $\psi_{j,k}(t) := 2^{-j/2}\,\psi\big(2^{-j}t - k\big)$, $\phi_{j,k}(t) := 2^{-j}\,\phi\big(2^{-j}t - k\big)$, $j, k \in \mathbb{Z}$ form an orthonormal basis, and we have the representation [3,4]

$$z(t) \;=\; \sum_k u_{j_0,k}\,\phi_{j_0,k}(t) \;+\; \sum_{j=-\infty}^{j_0}\sum_k w_{j,k}\,\psi_{j,k}(t), \tag{3}$$

with $u_{j,k} := \int z(t)\,\phi^*_{j,k}(t)\,dt$ and $w_{j,k} := \int z(t)\,\psi^*_{j,k}(t)\,dt$. 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 [3,4].

The DWT enjoys an enviable energy compaction property: the energy of many real-world signals compacts into just a few large wavelet coefficients, while white noise remains disbursed over a large number of small coefficients. This disparity can be exploited to distinguish signal from noise and has given rise to a number of powerful denoising techniques based on simple thresholding [4–6] that can suppress noise while preserving time-localized signal structures.

Wavelet denoising figures prominently in a number of recent advanced deconvolution algorithms [3, 7, 8]. All three methods have the same two basic steps in common:

Inversion: Compute the noisy estimate $\widehat{X}_0 = H^{-1}(f)\,Y(f)$. This inversion necessarily amplifies noise components at frequencies where $H(f)$ is small.

Regularization by Wavelet Denoising: Compute the DWT of $\widehat{x}_0$, and then threshold and invert the DWT to obtain the final signal estimate $\widetilde{x}$. Note that the Inversion step colors the white corrupting noise n; hence scale-dependent thresholds [6] should be employed.

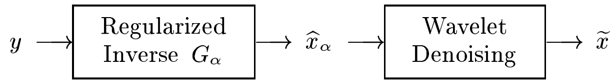

Figure 1: *Wavelet-based regularized deconvolution (WaRD): partially regularized inverse filtering following by wavelet denoising.*

4. IMPROVED WAVELET-BASED DECONVOLUTION WITH REGULARIZED INVERSE

The current Fourier/wavelet deconvolution approaches of [3, 7, 8] completely decouple the inversion and regularization processes. Unfortunately, when the system h is very ill-conditioned (or simply non-invertible), any attempt at inversion will amplify the corrupting noise to an extent that it will obliterate the desired signal. No amount of wavelet denoising can rescue us in this case.

Note, however, the sensitivity of the inversion process to regularization. A minute amount of regularization (small α in (2)) can lead to a huge reduction in the degree of noise amplification - all this at the expense of only a slight increase in the signal distortion. This realization motivates us to replace the Inversion step of the algorithms of [3, 7, 8] with a *Regularized* Inversion step (see Figure 1).[1] We call the resulting algorithm *wavelet-domain regularized deconvolution* (WaRD).

But how to pick the right value for the regularization parameter α? The tradeoff is clear: On one hand, since regularization smears non-stationary signal features like edges and ridges, we would prefer α as small as possible. On the other hand, large α prevents excessive noise amplification during inversion which aids the wavelet denoising.

To be more precise, we will determine the *optimal* regularization parameter for the WaRD system by minimizing the overall MSE. The deconvolution MSE consists of the signal distortion due to Fourier-domain regularization and the error due to wavelet-domain denoising:

$$\begin{aligned} \mathrm{MSE}(\alpha) \;&=\; \int [1 - G_\alpha(f)\,H(f)]^2\,P_x(f)\,df \\ &\quad + \sum_{j,k} \min\big(|\theta_{j,k}(\widehat{x}_\alpha)|^2,\, \sigma_j^2(\alpha)\big)\,. \end{aligned} \tag{4}$$

Here $\theta_{j,k}(\widehat{x}_\alpha)$ denotes the wavelet coefficients of $\widehat{x}_\alpha$, $\sigma_j^2(\alpha)$ is the variance of the wavelet-domain noise at scale j, and $P_x(f) = |X(f)|^2$.

The first term in $\mathrm{MSE}(\alpha)$ is an estimate of the distortion in the input signal due to the regularized Fourier-domain inverse [9]. This distortion is an increasing function of α. The second term is an estimate of the error due to ideal wavelet domain hard thresholding [5]. Ideal thresholding consists of keeping 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. (Ideal thresholding assumes that the signals under consideration are known.) This error is a decreasing function of α. The optimal regularization parameter, denoted by $\alpha^\star$, corresponds to the minimum of $\mathrm{MSE}(\alpha)$. Note that $\alpha^\star$ depends on the signal, system, and noise power.

The existing Fourier/wavelet deconvolution algorithms of [3, 7, 8] can be interpreted as special cases of WaRD with $\alpha = 0$.

[1] Note that the Fourier-domain-regularized inverse (2) was derived for WSS signals x only. With non-stationary x, we replace P_x by the time-averaged spectrum of x. In practice, we set $P_x(f) = |X(f)|^2$.

However, as mentioned earlier, these methods are in general not applicable when h is not invertible. Even when h is invertible, WaRD will outperform these methods, since the value $\alpha = 0$ is included in the search-space for the optimal $\alpha^\star$.

The computational cost of deconvolving an M-point signal will be dominated by the $O(M \log M)$ cost of the FFT inverse-filter implementation. The second, wavelet denoising step consumes only $O(M)$ computations.

The wavelet denoising step of the WaRD algorithm can be extended in several ways. First, since the standard DWT is not shift-invariant, shifts of y will result in different estimates $\widetilde{x}$. Employing a redundant, shift-invariant DWT will both yield a shift-invariant algorithm as well as improve the denoising performance substantially [4], all at no significant increase in the overall computational cost. The recently proposed complex and "almost shift-invariant" DWT of [10] would yield similar results at a reduced computational cost. Finally, instead of a threshold, we can apply a Wiener filter to the *wavelet* coefficients [11, 12].[2] Such processing has been shown to outperform simple thresholding for denoising finite samples of data.

Finally, note that WaRD extends trivially to higher dimensions using the appropriate Fourier and wavelet transforms.

5. EXAMPLES

To illustrate the performance of the WaRD algorithm, we will perform simulations in 1-d and 2-d.

We first compare the WaRD with other methods for the 1-d deconvolution problem presented in the Introduction. In order to observe the behavior of each method for both smooth and edgy regions, we take for x a concatenation of Donoho's *Blocks* and *Heavisine* signals [13] (normalized to be zero mean and unit energy). We employ Daubechies length-8 wavelets throughout. Figure 2(a) depicts the signal. For the system, we take the example of [3, pp. 459]

$$H(f) = \begin{cases} 1, & |f| \in [0, 0.25] \\ 2 - 4\,|f|, & |f| \in (0.25, 0.5] \end{cases} \qquad (5)$$

with frequency f normalized to $(-0.5, 0.5]$ (see Figure 2(b)). The corrupting noise variance was $\sigma^2 = 4 \times 10^{-6}$. Figure 2(c) plots the blurred, noisy signal.

The Wiener filter estimate (Figure 2(d)) was implemented using $\alpha = 1$ and $P_x(f) = |X(f)|^2$ in (2). The Wiener filter bases its deconvolution on the signal-to-noise ratio at each frequency. However, this is inappropriate for non-stationary signals, since their frequency content changes with time. The deconvolution suffers in response.

The methods of [7, 8] fail in this case, due to the null in the frequency response of the system at $H(0.5) = 0$.

Since $H(f)$ decays to zero at $f = 0.5$ so slowly, the *wavelet packet deconvolution method* of [3, pp. 458–461] remains applicable.[3] This method adapts a wavelet packet basis to the colored noise in the inverted data $\widehat{X}(f) = H^{-1}(f)Y(f)$. The frequency splits of the wavelet packet basis for H are shown with dashed lines in Figure 2(b). The denoising step was implemented by hard-thresholding the coefficients of a shift-invariant DWT using the best wavelet packet basis. This algorithm outperforms the methods of [7, 8] typically, as well as the standard Wiener filter in this case (see Figure 2(e)).

Figure 2(f) plots the WaRD obtained using $\alpha^\star = 0.06$. Wavelet-domain Wiener filtering was applied to the coefficients of a shift-invariant DWT for the denoising stage. The WaRD outperforms the other algorithms in terms of both visual quality and MSE performance.

Next, we consider image restoration using WaRD (ReWaRD). The input x is the 256×256 Lena image (normalized to zero mean and unit energy) and the discrete-time system response h is a 2-d, 4-point smoother $[1\ \ 1\ \ 1\ \ 1]^T\,[1\ \ 1\ \ 1\ \ 1]$. 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 noise variance was set to $\sigma^2 = 4 \times 10^{-7}$. Figure 3 illustrates the desired x, the observed y, the Wiener filter estimate $\widehat{x}_1$, and the WaRD estimate for $\alpha^\star = 0.27$. The methods of [3, 7, 8] are not applicable in this situation, due to the many zeros in $H(f_x, f_y)$. (The Wiener filter outperformed the wavelet packet method in this case.) WaRD clearly outperforms Wiener filtering in both visual quality and MSE.

6. CONCLUSIONS

In this paper, we have proposed an efficient multi-scale deconvolution algorithm that optimally combines Fourier-domain regularized inversion and wavelet-domain denoising. For non-stationary signals, the WaRD outperforms the LTI Wiener filter and other wavelet-based deconvolution algorithms in terms of both visual quality and MSE performance. Furthermore, it continues to provide a good estimate of the original signal even when the system response is ill-conditioned. All of this in an algorithm of computational complexity no greater than an FFT.

For a given problem setup, the optimal value of the regularization parameter α depends on the signal, the system response, and the noise level. Fortunately, final performance was observed to be quite insensitive to the exact value, as long as we choose alpha sufficiently positive. As a guide, in simulations spanning many real-world images and convolution systems, $\alpha^\star$ was almost always lay in the range $[0.2, 0.3]$. In the 1-d example above, the optimal α turned out to be small ($\alpha^\star = 0.06$) because the test signal compacted very well in the wavelet domain. However, choosing a $\alpha^\star$ in the range $[0.2, 0.3]$ gave near-optimal results.

There are several avenues for future WaRD related research. We are developing methods to estimate the WaRD MSE, and thus the optimal $\alpha^\star$, without prior knowledge of the signal and noise power. Further, we are investigating the gains possible using a "best basis" adapted to the *signal* instead of the noise as suggested in [3, pp. 458–461].

7. REFERENCES

[1] J. G. Proakis, *Digital Communications*. McGraw-Hill, 1995.

[2] A. K. Jain, *Fundamentals of Digital Image Processing*. Prentice-Hall, 1989.

[3] S. Mallat, *A Wavelet Tour of Signal Processing*. Academic Press, 1998.

[4] C. S. Burrus, R. A. Gopinath, and H. Guo, *Introduction to Wavelets and the Wavelet Transform*. Prentice-Hall, 1998.

[2] Do not confuse wavelet-domain Wiener filtering with the Fourier-domain Wiener filtering discussed above.

[3] Even though $H(f)$ is not invertible, the wavelet packet deconvolution method does not fail, because the location and the order of the zero of H are such that only a few, high-frequency signal wavelet coefficients are obliterated by the infinite noise amplification at $f = 0.5$. This method will fail when $H(f)$ has zeros of arbitrary location and order, however.

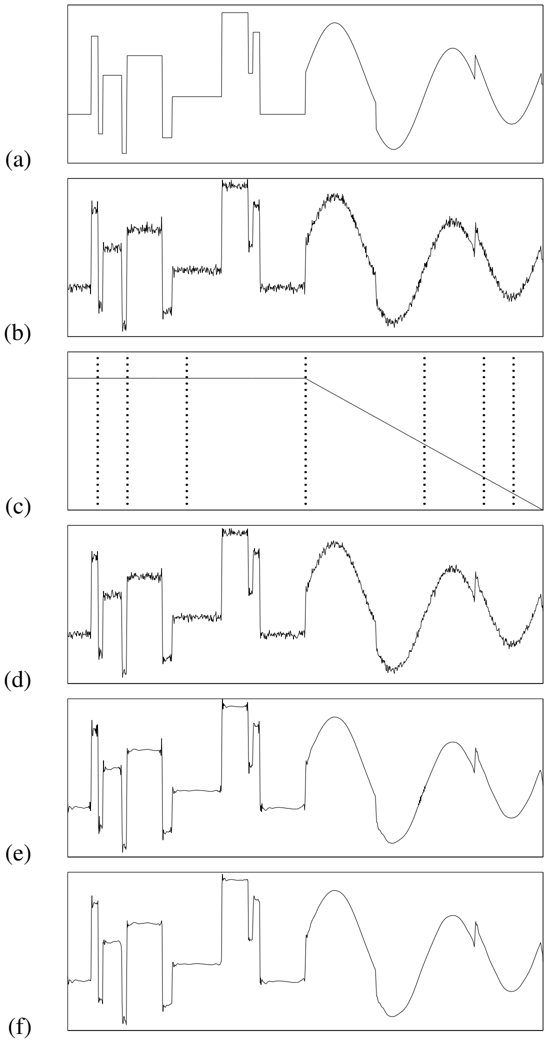

Figure 2: *(a) Test signal x ($M = 1024$). (b) Observed signal y (MSE=0.077). (c) System frequency response $H(f)$. (Wavelet packet tree frequency splits for estimate (e) shown with dashed vertical lines.) (d) Wiener filter estimate (MSE=0.069). (e) "Wavelet packet" estimate [3, pp. 458–461] (MSE=0.052). (f) WaRD with $\alpha^\star = 0.06$, (MSE=0.04).*

Figure 3: *Upper left: Lena image x (256×256). Upper right: Observed image y (MSE=0.457). Lower left: Wiener filter estimate (MSE=0.199). Lower right: WaRD with $\alpha = 0.27$ (MSE=0.182).*

[5] 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.

[6] I. M. Johnstone and B. W. Silverman, "Wavelet threshold estimators for data with correlated noise," *J. Royal Stat. Soc. B*, no. 59, pp. 319–351, 1997.

[7] D. L. Donoho, "Nonlinear solution of linear inverse problems by Wavelet-Vaguellete Decomposition," *Appl. Comp. Harm. Anal.*, vol. 2, pp. 101–126, 1995.

[8] R. Nowak, "A fast Wavelet Vaguelette algorithm for discrete LSI problems," *IEEE Sig. Proc. Lett.*, Submitted Aug. 1997.

[9] N. P. Galatsanos and A. K. Katsaggelos, "Methods for hoosing the regularization parameter and estimating the noise variance in image processing and their relation," *IEEE Trans. on Image Proc.*, vol. 1, pp. 322–336, Jul. 1992.

[10] N. Kingsbury, "The dual-tree complex Wavelet Transform: A new technique for shift invariance and directional filters," in *IEEE DSP Workshop*, (Bryce Canyon, UT), Aug. 9-12 1998.

[11] S. Ghael, A. M. Sayeed, and R. G. Baraniuk, "Improved wavelet denoising via empirical Wiener filtering," in *Proc. of SPIE*, vol. 3169, (San Diego), Jul. 1997.

[12] H.Choi and R. G. Baraniuk, "Analysis for Wavelet Domain Wiener Filters," in *Proc. IEEE Conf. on Time-Frequency and Time-Scale Analysis*, (Pittsburgh), Oct. 1998.

[13] D. L. Donoho and I. M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," *J. Am. Stat. Assoc.*, vol. 90, pp. 1200–1224, 1995.