IMAGE DENOISING USING MULTIPLE COMPACTION DOMAINS

Prakash Ishwar, Krishna Ratakonda, Pierre Moulin and Narendra Ahuja

University of Illinois at Urbana–Champaign
Beckman Institute and ECE Department
405 N. Mathews Avenue
Urbana, IL 61801.
{ishwar,ratakond}@uiuc.edu

ABSTRACT

We present a novel framework for denoising signals from their compact representation in multiple domains. Each domain captures, uniquely, certain signal characteristics better than others. We define *confidence sets* around data in each domain and find sparse estimates that lie in the intersection of these sets, using a POCS algorithm. Simulations demonstrate the superior nature of the reconstruction (both in terms of mean–square error and perceptual quality) in comparison to the adaptive Wiener filter.

1. INTRODUCTION

We consider the problem of estimating a signal corrupted with zero–mean additive white Gaussian noise (AWGN) of a known variance σ^2. The Wiener filter is optimal in minimizing the mean–square error under suitable assumptions on the stationarity of the signal statistics. Locally, such assumptions are reasonable, as in the adaptive realization [1] of the Wiener filter whose performance is among the best known till date.

Recently, there has been renewed interest in various threshold–based denoising methods, in linear unitary transform domains where the signal has a sparse representation. Towards this end, wavelets are known to possess this compaction property for a wide class of real–life signals. Thresholding in the wavelet domain has been shown to be asymptotically nearly optimal in a minimax mean-square error (MSE) sense over a variety of smoothness spaces [2]. Donoho's hard thresholding, for example, "kills" transform coefficients smaller than the universal threshold $\sigma\sqrt{2\ln(N)}$ (where N is the data size), by setting them to zero and retains others unaltered. The philosophy of this approach is that small coefficients are very likely due to noise, the threshold representing an optimal tradeoff in terms of the "risk" incurred in destroying a significant signal coefficient (when the threshold is too large) versus retaining excessive noise power (when the threshold is small). Despite these asymptotic optimality properties, hard thresholding produces prominent artifacts when applied to images corrupted by moderate to high amounts of noise.

We now dwell briefly on the central theme of the paper, an idea that is expounded in the sections to follow. Assuming for the moment that a universally thresholded noisy signal is a "reasonably reliable" signal estimate, a natural question is the choice of the unitary transform in which the signal representation is sparse. If one were to restrict oneself to a class of structured linear unitary transforms, such as for example, the class of all orthonormal wavelet transforms, there would in general exist several domains in each of which the signal energy is "sufficiently compacted" in a relatively few, large coefficients and there is really no strong reason to prefer one domain to another in terms of the MSE performance. There is clearly an information overlap in these equivalent representations but at the same time, each domain captures, uniquely, certain signal characteristics better than others, an observation that is also corroborated in our study. We therefore expect that, "intelligent" combination of estimates in different transform domains should improve performance. We restrict our investigation to combining information from Donoho's hard–thresholded estimates, but the principle is by no means restrictive to a specific denoising scheme or to a specific class of signal representation domains. The key idea behind our scheme is to define *confidence sets* around the data in each transform domain. We then seek a signal that lies in the intersection of these sets and, ideally, is closer to the signal than the estimate in any single domain.

We have recently learned about another method for denoising using two wavelet domains [3]. There, a second domain was used to obtain a (pilot) signal estimate from which to compute signal statistics to design a wavelet–domain Wiener filter.

2. OVERVIEW OF THE MULTIPLE–DOMAIN DENOISING ALGORITHM

To start with, we need to identify the wavelet filters which would be best for complementary processing. The information lost in any domain by thresholding is represented by small valued coefficients in that domain. One would want the information captured by relatively small coefficients in one domain to be represented by large valued coefficients (that would not be destroyed by thresholding) in others while simultaneously desiring a compact signal representation in all domains. We leave the problem of optimal transform class selection for future research. One possible heuristic for selecting the wavelet domains is to use filters that generate wavelets having different degrees of smoothness. In our simulations, we use a family of orthonor-

mal, compactly-supported wavelets with extremal phase and minimum number of vanishing moments for the given support length, due to Daubechies [4]. This family was found to possess the desired properties for most 2-D signals such as *Lena* and *Sailboat* that we worked with.

In each domain we define a *confidence tube* of *radius* δ as the set of signals with coordinates s_i satisfying $|s_i - d_i| \leq \delta_i$, where d_i are the transform data and $\delta_i = \delta > 0$ if $|d_i| < \lambda = \sigma\sqrt{2\ln(N)}$ and $\delta_i = 0$ otherwise. (See Fig. 1). By defining $\delta_i = 0$ for large coefficients, we force the estimates to agree with "reliable" components of the data.
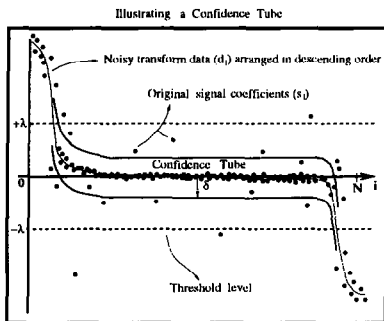

Figure 1: 2-D representation of a *confidence tube* with radius δ centered around noisy transform data (arranged in descending order).

The confidence tube in each domain is both closed and convex. This suggests the use of successive projections of an initial signal estimate onto confidence tubes in multiple domains, as a means of extracting a signal that retains characteristics of all tubes while being close to the initial signal estimate. That, successive projections onto closed convex sets starting from an arbitrary initialization will converge, is a result well known in the theory of projection onto convex sets (POCS) [5].

POCS seeks a point that lies in the intersection of several closed convex sets, the initial point being important only so far as to affect the dynamics of convergence. In the problem at hand however, the initial point is crucial, since, the intersection set contains many undesirable signals, including, for example, the noisy data. It can be shown that the intersection set also contains much sparser signals. We would like the final iterate to inherit the characteristics of "reasonable" estimates in every domain while being closer to the clean signal. (Other convex constraints that might help characterize the signal can be easily incorporated into this framework. See below.). A reasonable choice for the initial point is thus the hard-thresholded signal estimate in any domain. An equivalent choice is the zero signal (the proof is elementary).

3. NUMERICAL RESULTS

Fig. 2 compares the PSNR ($10\log_{10}(\frac{255^2}{MSE})$) performance for *Lena* as a function of tube radius for noise with s.d. 10 (with the noisy data in the image domain quantized to integers in the range 0–255). Here, Daub. 3,4 refers to the Daubechies filters with 3 and 4 vanishing moments. The limited spatial-

intensity range, an additional convex constraint popular in POCS based image restoration literature, was also incorporated in our simulations. Observe that reconstructions using more domains (if appropriately chosen), do consistently better for most tube radii. Also observe that the peak is consistently attained at around 2σ. The peak performance using four domains was better by 0.22 dB than MATLAB's spatial, adaptive, Wiener filter reconstruction using a window of size 3×3. Several window sizes were tried and the one that produced the smallest MSE was chosen.
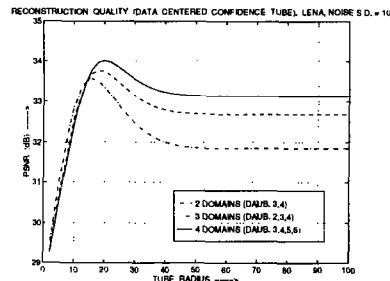

Figure 2: Comparison of performance on *Lena* using data-centered confidence tubes, $\sigma = 10$. Dot-dashed line : two domains (Daub. 3,4). Dashed line : three domains (Daub. 2,3,4). Solid line : four domains (Daub. 3,4,5,6).

In Fig. 3, we compare the perceptual quality of image reconstruction by different methods for *Lena* corrupted with AWGN of s.d 15. Fig. 3(a) and (b) show the original and noisy *Lena* image, respectively. Fig. 3(c) shows the result of wavelet (Daub. 3) hard thresholding using Donoho's threshold. Notice the ringing artifacts near edges that arise due to zeroing out of significant edge information present in small valued coefficients. Fig. 3(d) shows MATLAB's adaptive Wiener filter reconstruction (in the spatial domain) using a window of optimal size, 3×3. Notice that this reconstruction contains significant residual noise, and the image still looks grainy. Fig. 3(e) shows the multiple-domain reconstruction using two domains (Daub. 3,4), and Fig. 3(f) shows a 5-domain reconstruction (Daub. 3,4,5,6,7) for a tube radius of 2σ. Observe that detail features in the texture of the hat and hair and the lip shape are better captured in the 5-domain reconstruction than the other methods. The image in Fig. 3(c) was obtained by performing hard thresholding in each of these five domains and choosing the reconstruction that produced the smallest MSE (although there was only marginal variation in PSNR in different domains).

Further improvements in performance are attained by averaging out the final iterates corresponding to initial points (the hard-thresholded estimates) from each domain. Except for results of Fig. 2, whenever we use two domains, we perform this additional averaging of final iterates. This is quite different from taking a simple average of the hard-thresholded estimates in different wavelet domains. For *Lena* corrupted by noise of s.d. 10, the PSNR of the best hard-thresholded estimate among the two domains was 30.41 dB, a simple averaging of these estimates had a PSNR = 31.42 dB, while the 2-domain reconstruction with averaging of final estimates had a PSNR = 33.68 dB for a tube

radius of 2σ. While convergence of the algorithm is typically attained in eight iterations, we have used ten iterations in all our simulations, where, a complete sequence of successive projections onto convex tubes in multiple domains, together with the imposition of spatial, signal intensity convex constraints between these projections, constitutes one iteration. In the following section we examine merits related to centering the confidence tube around data as against the estimate (the popular practice in POCS literature). We also examine some sparsity properties of the proposed scheme. Towards this end, we restrict our discussion to two wavelet domains (Daub. 3,4).
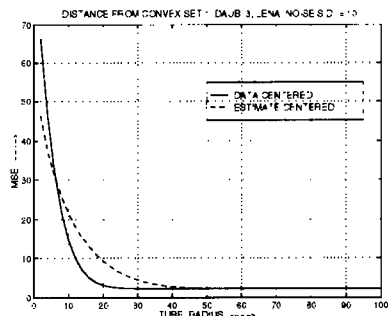

Figure 4: Distance from convex set in domain 1 (Daub. 3), for *Lena*, $\sigma = 10$, using two domains (Daub. 3,4). Solid line : data–centered. Dashed–dot line : estimate–centered.

4. DISCUSSION

In hypothesis testing literature, *confidence intervals* are classically centered around estimates. Our approach of using data-centered estimates is a significant departure from convention and needs justification.

In order to simplify the discussion, we shall assume a uniform confidence tube centered around *all* noisy coefficients, that is, $\delta_i = \delta, \forall i$[1]. The apparent distribution of the original signal values around the noisy data is i.i.d. $N(0,\sigma^2)$. Thus at each noisy coefficient, the original signal coefficient is expected to be found (with a high probability) within an interval of radius 2σ. It can be shown that the expected Euclidean distance of the original signal from confidence tubes centered around data is independent of the signal values. For confidence tubes centered around thresholded estimates however, this expected distance is data–dependent. For highly sparse signals, estimate–centered confidence intervals are the better choice. However, there exist classes of signals for which the expected Euclidean distance from the data–centered set is smaller than that from an estimate–centered one. That, for many real–life images, a data-centered tube is closer to the original image than an estimate-centered one is corroborated in our simulations, see Fig. 4. Fig. 5 shows the mean–square distances of successive projections from the original signal for a tube radius that was best (in terms of MSE) for each of the convex sets (the data

and estimate–centered tubes). Observe that successive distances in each case diminish with successive projections as expected, but the data- centered reconstruction gets closer to the original signal at each stage. Fig. 6 compares the PSNR performance of either set as a function of tube radius. The curves for each set behave quite differently, the data–centered estimates performing consistently better for all tube radii.

An interesting result is the following : if the initial estimate is the zero signal, its first projection onto a data centered set is *equivalent* to performing a *soft thresholding* operation on the noisy data with a threshold equal to the tube radius, δ. The soft thresholding operation is given by $sgn(d)\max(0,|d| - \delta)$ where d is the noisy data.

The final iterate has a sparse representation as measured by the histogram of its coefficients. Fig. 7 compares the histograms of transform coefficients for the original and noisy images, the hard–thresholded estimate, and the 2 domain, data centered reconstruction. Observe that the multiple–domain, data–centered reconstruction is sparse and is closer to the original histogram than is the hard–thresholded reconstruction.

5. CONCLUDING REMARKS

We have demonstrated a novel method of combining information from multiple signal representations, to significantly improve a baseline denoising scheme such as hard thresholding that produces oversmoothing and ringing artifacts near edges when applied to typical images. The gains in PSNR are typically of the order of 2 to 3 dB for typical real–life images. Similar gains (not discussed here) were also obtained for smaller and larger noise variances. The multiple–domain approach has opened up possibilities of extracting "useful" information from several (maybe poor) estimates of the signal to improve performance by huge margins. These principles are not restricted to corruption by AWGN. There are several avenues for future work. An important issue is the optimal selection of representation domains. Also needed is a rigorous analysis of the method and ways to incorporate statistical knowledge of the signal/noise into the framework of confidence sets.

6. REFERENCES

[1] D. T. Kuan, A. A. Sawchuk, T. C. Strand and P. Chavel, "Adaptive Noise Smoothening Filter for Images with Signal-dependent Noise," *IEEE Trans. on Patt. Anal. and Mach. Intell.*, Vol. 2, pp 165–178, 1980.

[2] D. L. Donoho and I. Johnstone, "Ideal Spatial Adaptation by Wavelet Shrinkage," *Biometrika*, Vol. 81, pp. 425–455, 1994.

[3] S. P. Ghael, A. M. Sayeed and R. G. Baraniuk "Improved Wavelet Denoising via Empirical Wiener Filtering," *Proc. SPIE, Mathematical Imaging*, July 1997.

[4] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, 1992.

[5] H. Stark and M. I. Sezan, "Image Processing Using Projection Methods," *Real–Time Optical Information Processing*, pp. 185- 232, Academic Press, 1994.

---

[1]However, it has been our experience that estimation results are then somewhat worse than those presented in Sec. 3.

(a)

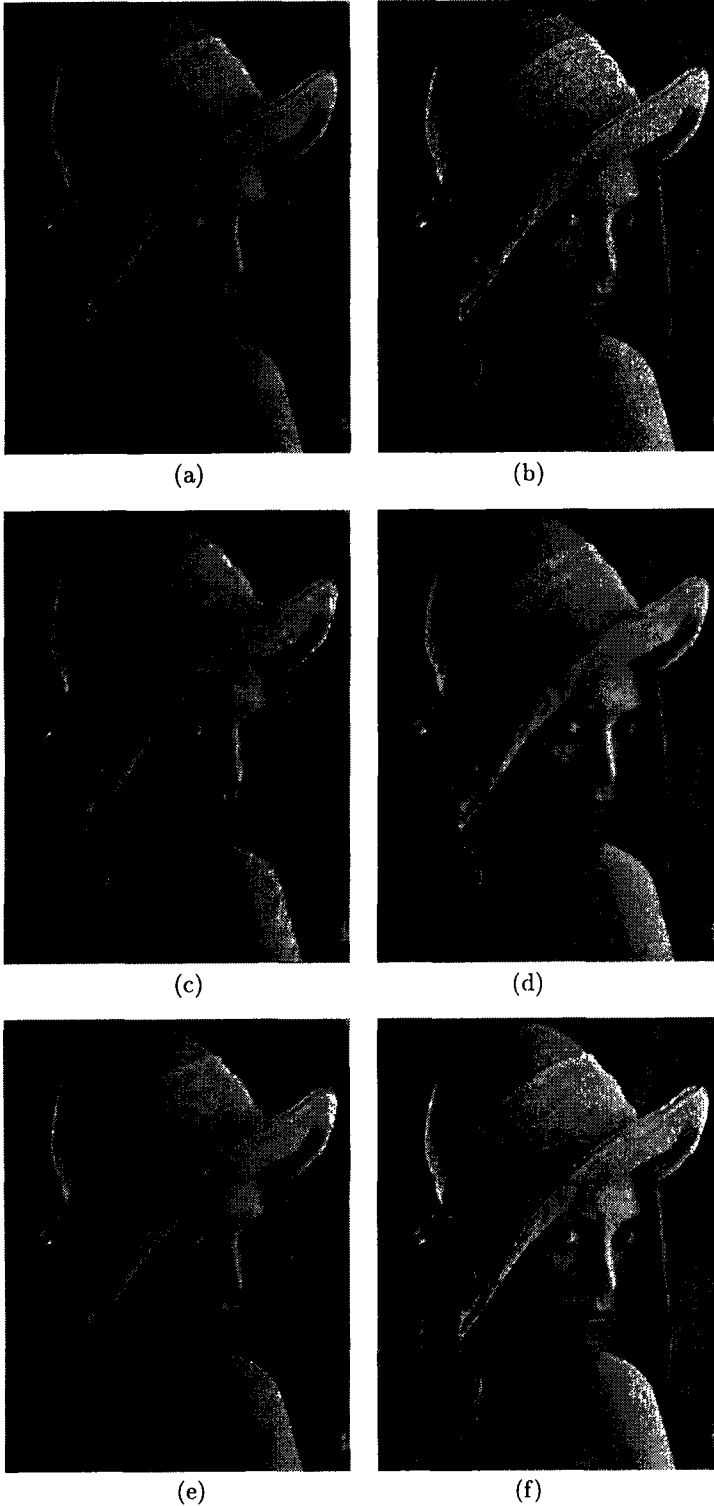

(b)

(c)

(d)

(e)

(f)

Figure 3: Comparison of different denoising schemes. (a) Original Image, (b) Noisy *Lena*, $\sigma = 15$, PSNR = 24.61 dB; (c) Donoho's wavelet thresholding (Daub. 3), PSNR=28.54 dB; (d) MATLAB's adaptive Wiener filter (3×3 window), PSNR=31.24 dB; (e) 2-domain reconstruction (Daub. 3,4), PSNR=31.38 dB; (f) 5-domain reconstruction (Daub. 3,4,5,6,7), PSNR=31.76 dB
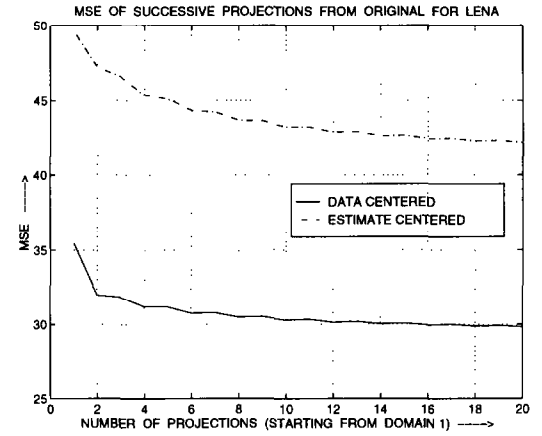

Figure 5: MSE of successive projections for *Lena*, $\sigma = 10$, optimal tube radii, two domains (Daub. 3,4).
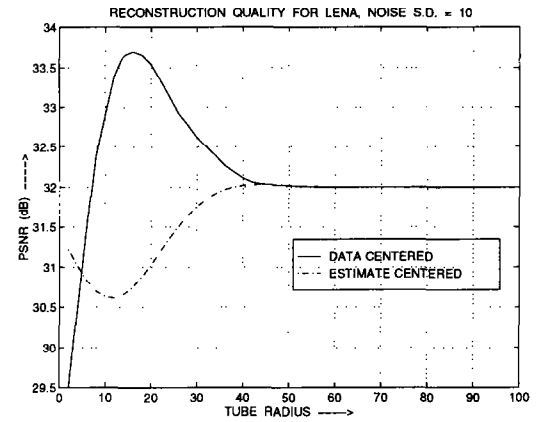

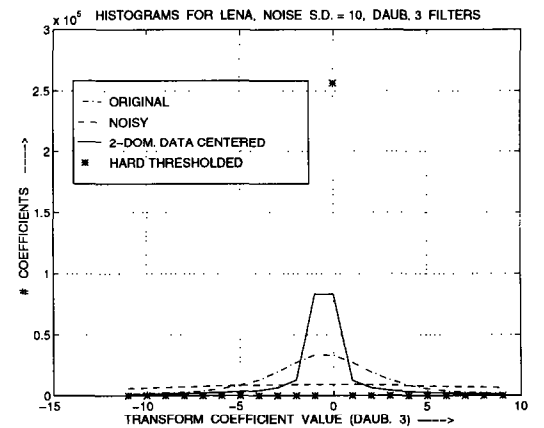Figure 6: PSNR performance for *Lena*, two domains (Daub. 3,4).

Figure 7: Comparison of histograms of transform coefficients (Daub. 3) for *Lena*. Dashed–dot line : original. Dashed line : noisy($\sigma = 10$). Solid line : 2–domain, data–centered. Stars : hard–thresholded.