

SPATIALLY ADAPTIVE IMAGE DENOISING UNDER OVERCOMPLETE EXPANSION

Xin Li and Michael T. Orchard

Department of Electrical Engineering
Princeton University
Email: lixin, orchard@ee.princeton.edu

ABSTRACT

This paper presents a novel wavelet-based image denoising algorithm under overcomplete expansion. In order to optimize the denoising performance, we make a systematic study of both signal and noise characteristics under overcomplete expansion. Highband coefficients are viewed as the mixture of non-edge class and edge class observing different probability models. Based on improved statistical modeling of wavelet coefficients, we derive optimal MMSE estimation strategies to suppress noise for both non-edge and edge coefficients. We have achieved fairly better objective performance than most recently-published wavelet denoising schemes.

1. INTRODUCTION

There has been growing interest of wavelet-based denoising schemes for removing additive white Gaussian noise from corrupted images recently. Such popularity is mainly due to that wavelet provides an appropriate basis for separating noise signal from image signal. Important features in images(e.g. edges) are usually represented by large coefficients, thus facilitating their separation from noise components which usually correspond to small coefficients in wavelet domain. Even ad-hoc thresholding strategy [1] generates satisfactory denoising result and often outperforms other class of denoising schemes such as nonlinear filtering or nonlinear diffusion in terms of objective performance (MSE or PSNR).

When noise characteristics is additive white Gaussian, image denoising problem is closely related to lossy image coding problem studied in [7] because they both heavily rely on statistical models employed in wavelet domain. Since lossy image coding has been extensively studied in image processing community, it is natural to make use of modeling maturity in lossy image coding to enhance the performance of wavelet-based image denoising algorithms. Indeed, recent works e.g. [2, 3, 4] follow this trend and have demonstrated impressive improvements over ad-hoc thresholding strategy. On the other hand, due to different goals between

image denoising and image coding, appropriate wavelet basis to expand image signal becomes different. Redundant representation or overcomplete expansion is obviously inefficient for coding purpose. However, it has been shown in [5, 4] that redundant representation is substantially superior to non-redundant representation for denoising purpose. [5] summarizes the benefits of using redundant representation as improvement in approximation and suppression of Gibbs phenomena.

We present a new image denoising scheme using overcomplete expansion in this paper. The overcomplete expansion is obtained in a similar fashion as Translation-Invariant (TI) denoising [5] by shifting original signal. But unlike [5] processing shifted versions separately, we propose to put them together for the reason of achieving more accurate statistical models for signal components. The highband coefficients are viewed as the mixture of non-edge coefficients observing a zero-mean model and edge coefficients observing a biased-mean model. The optimal noise suppression strategies(MMSE estimation) are derived for both edge and non-edge coefficients. For edge coefficients whose true bias is not available, an approximated solution based on Least-Square estimation and directional filtering is adopted. The new denoising algorithm appears to give smaller MSE results than almost all recently-published results in the literature of wavelet denoising. The rest of this paper is organized as following. Section 2 studies overcomplete expansion and its effect on signal and noise characteristics. Section 3 derives the optimal MMSE estimation for non-edge and edge coefficients under overcomplete expansion separately. Section 4 includes our experiment results for the new denoising algorithm.

2. OVERCOMPLETE EXPANSION

The preference of overcomplete expansion over conventional orthogonal wavelet originated from Coifman and Donoho's Translation-Invariant(TI) denoising [5]. In 1D case, translation invariance is achieved by shifting the input signal X by $K - 1$ indices to obtain K redundant representations where K is the decimation ratio in orthogonal wavelet transform.

Then denoising operation is applied to all shifted versions and the average of K denoised signal is taken as the final result $\hat{s}$. The whole procedure can be written as:

$$\hat{s} = \frac{1}{K} \sum_{k=0}^{K-1} Unshift_k[Denoise[Shift_k[X]]] \quad (1)$$

where *Unshift* is just the inverse operation of *Shift*. We also obtain overcomplete expansion by shifting the input signal by $K - 1$ indices to generate K redundant representations. However, we look at the overcomplete expansion from point of view of its effect on signal and noise characteristics. The new perspective not only helps us better understand the preference of overcomplete expansion over orthogonal expansion in denoising scenario but also motivates us to come up with a better strategy for noise suppression.

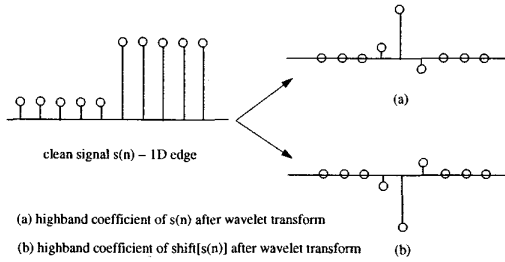


Fig. 1. Overcomplete expansion of a 1D step-edge signal.

Let us use a 1D step edge s_n as an example to explain the advantage of using overcomplete expansion. Similar arguments also hold for 2D signal. Since wavelet transform is linear, the effect of noise on signal can still be treated as additive in wavelet domain. So we can separately study signal and noise components. On signal part, figure 1 shows the highband coefficients of all shifted version of s_n . Our interest actually only lies in large coefficients(spikes) since they carry the information about transitions. When step edge is shifted by one sample, the spike changes its sign but still remains as a spike. On noise part, one obvious fact is its correlated structure(e.g. version (a) and (b) are correlated). However, the noise is still independent of s_n and its power is typically much lower than that of signal component(spikes). Consequently, the probability of both spikes disappearing is much less then the probability of either one disappearing, which explains why signal is better preserved under overcomplete expansion. Averaging over all shifted versions is simply a convenient solution without resolving the uncertainty of in which version the spike is gone.

Theoretically jointly processing of noisy coefficients is necessary for the optimal estimation in the case of correlated-noise. In practice independently processing each coefficient

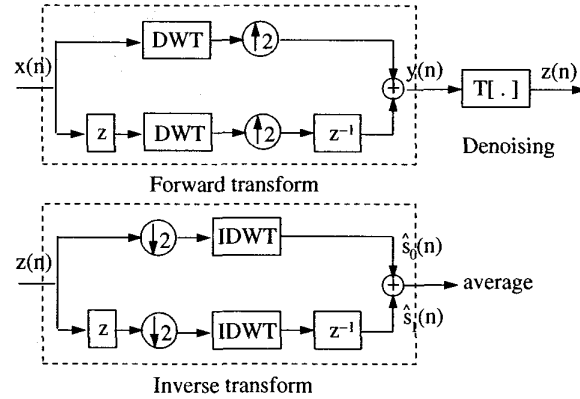


Fig. 2. Diagram of our new denoising approach.

is preferred due to its simplicity. This is dual to the preference of scalar quantization over vector quantization in image coding. When we are constrained to a scalar version of denoising scheme, the correlated structure of noise does not impede the derivation of optimal estimation strategy as we shall detail in the following section. The MMSE estimation actually only needs the assumption of zero-mean and independence from the signal for noise characteristics.

Based on the above observations, we propose to process all shifted versions jointly rather than independently like [5, 4]. In spite of correlated noise structure, the task of accurately statistical modeling of signal components is greatly facilitated. For example, at the highest resolution, the overcomplete expansion is equivalent to the result of wavelet filtering without decimation. Since wavelet filtering is linear, many nice properties of edges such as geometric constraint [7] are well preserved in wavelet domain. And statistical modeling of edge coefficients could benefit from those properties. To clarify the distinction between our approach and former works [5, 4], figure 2 shows the diagram of our approach in 1D case. Generalization into 2D case is straightforward. Redundancy ratio in 2D case would be 4 instead of 2.

3. MMSE ESTIMATION

Since we process each coefficient independently, the efficiency of noise suppression strategy mainly depends on the accuracy of statistical models describing signal and noise characteristics under overcomplete expansion. All the derivation of MMSE estimation below will assume that noise is zero-mean and independent of signal but might be correlated. Similarly as [7], highband coefficients are viewed as the mixture of two classes: edge coefficient observing a biased-mean model whose biased mean depends on its lo-

cation across the edge orientation and non-edge coefficient observing a zero-mean model with spatially-varying variance. From now on, we use index n to denote spatial coordinate and subscript w for noise component.

3.1. Non-edge Coefficients

In smooth areas, we model wavelet coefficients X_n with $N(0, \sigma_n^2)$ and assume variance field σ_n^2 slowly varies, reflecting the spatial clustering energy within each band. The Maximum Likelihood (ML) estimation of σ_n^2 is given by [2]:

$$\hat{\sigma}_n^2 = \max[0, \frac{1}{|N_n|} \sum_{Y_k \in N_n} Y_k^2 - \sigma_w^2] \quad (2)$$

where local neighbor N_n is a square window centered at Y_n . Its size $|N_n|$ can be adaptively chosen across different resolution and orientation.

With the knowledge of σ_n^2 and σ_w^2 , it is not difficult to obtain the optimal linear MMSE estimation of X_n :

$$\hat{X}_n = aY_n, a = \frac{\sigma_n^2}{\sigma_n^2 + \sigma_w^2} \quad (3)$$

and the minimal MSE achieved by (3) is

$$MMSE = \frac{\sigma_n^2 \sigma_w^2}{\sigma_n^2 + \sigma_w^2} \quad (4)$$

It should be noted that correlated structure of noise does not affect the above derivation because it only employs the assumption of zero-mean and independence of signal. Putting all shifted versions together facilitates the estimation of signal characteristics σ_n^2 than treating them separately because they all contain the information about the same signal. Improved modeling of signal components directly leads to better objective performance of estimation strategy.

3.2. Edge Coefficients

We have found that biased-mean models are more suitable for edge coefficients in the framework of lossy image coding due to smoothness constraint along the edge orientation [7]. Actually the smoothness constraint is even better preserved under overcomplete expansion than orthogonal expansion as we mentioned at the end of section 2. Therefore we should be able to more accurately model edge coefficients under overcomplete expansion and improve the efficiency of estimation strategy.

However, if we still use the estimator $\hat{X}_n = aY_n$, we find that the efficiency is not really improved. When X_n observes a biased-mean model, the MMSE estimation a is:

$$a = \frac{E[X_n^2]}{E[X_n^2] + \sigma_w^2} \quad (5)$$

and the minimal MSE is:

$$MMSE = \frac{E[X_n^2] \sigma_w^2}{E[X_n^2] + \sigma_w^2} \quad (6)$$

We note that the MMSE result given by (6) is lower bounded by (4) because $E[X_n^2] = \sigma_n^2 + m_n^2 \gg \sigma_n^2$. Actually, for edge coefficients, we have:

$$E[X_n^2] \gg \sigma_w^2 \Rightarrow MMSE \approx \sigma_w^2 \quad (7)$$

It shows that noise has not been efficiently suppressed around edge area.

To enhance the performance of noise suppression for edge coefficients, it is possible to obtain an improved estimation of X_n by incorporating the biased mean m_n (temporarily assuming it is known by some oracle) into the estimation:

$$\hat{X}_n = aY_n + (1 - a)m_n, \quad (8)$$

It is easy to verify that (8) is an unbiased estimation of X_n ($E[\hat{X}_n] = E[X_n]$). Some simple derivation will show that the MMSE solution a goes back to (4) and minimal MSE reaches (4) again. Since $E[X_n^2] \gg \sigma_n^2$, (4) is dramatically smaller than (6), as we anticipate. Without the oracle about m_n , we propose the following approximated solution based on Least-Square estimation and directional filtering.

Firstly we estimate Y_n (actually its biased mean) from its eight nearest corrupted pixels (set N_n):

$$Y_n = \sum_{k \in N_n} \alpha_k Y_k \quad (9)$$

where $\vec{\alpha}$ is optimized through a standard Least-Square method inside a local square window centered at Y_n (including Y_n). Secondly, to smooth out the noise along the edge orientation, we take a simplified directional filtering approach:

$$\hat{X}_n = \frac{1}{|N_n^s|} \sum_{Y_k \in N_n^s} Y_k \quad (10)$$

where N_n^s is the subset of N_n :

$$N_n^s = \{Y_k \in N_n : |Y_k - Y_n| < 2\sigma\} \quad (11)$$

Intuitively, set N_n^s defined by (11) is an approximation of pixels which live along the same edge orientation as Y_n without explicitly estimating the orientation. In spite of the approximation, we find that directional filtering helps reduce MSE around strong edges.

4. SIMULATION RESULTS

We have compared our results to several current state-of-art schemes: 1) AWEQ (Adaptive Wiener EQ) [2] by M. Mihcak et al. of UIUC in ICASSP'99; 2) SSMM (Scale-Space

image	σ_w	10	12.5	15	20	25
Lena	AWEQ	23.7	-	37.6	52.1	67.5
	SSMM	21.5	-	36.0	-	-
	SANS	-	27.5	32.7	44.1	56.5
	Ours	20.9	26.7	32.5	44.1	56.6
Barbara	AWEQ	35.2	-	61.3	88.7	117.5
	SSMM	37.3	-	64.7	-	-
	SANS	-	43.5	56.0	83.1	112.2
	Ours	30.3	40.5	50.9	74.5	98.4

Table 1. MSE Performance comparison for Lena and Barbara

image	σ_w	10	12.5	15	20	25
Lena	9/7	21.3	28.8	34.0	47.1	61.0
	10/18	20.9	26.7	32.5	44.1	56.6
	28/28	20.6	26.3	31.9	43.3	55.4
Barbara	9/7	31.3	43.8	52.7	76.7	100.9
	10/18	30.3	40.5	50.9	74.5	98.4
	28/28	29.4	39.8	48.8	70.4	92.2

Table 2. MSE Performance comparison using different filters

Mixture-Modeling) [3] by J. Liu of UIUC in ICIP'99; 3) SANS(Spatially Adaptive Denoising) [4] by S.G.Chang of UC-Berkeley in ICIP'98. Among them SANS also used overcomplete expansion and achieved the best denoising results so far. Unlike other schemes using orthogonal filters, we employ a pair of biorthogonal 10/18 filters which have been used in lossy image coding [7]. The computational complexity of suppressing noise in wavelet domain is comparable to that of forward and inverse transforms.

Table 1 compares our denoising scheme with others in terms of MSE. For Lena image, our scheme gives noticeably smaller MSE than other schemes when the noise power(σ_w) is not strong. For Barbara image containing abundant strong edges, substantial gain has been achieved over former schemes. Figure 3 shows the comparison of denoised images using SANS [4] and ours. It can be observed that texture patterns around the scarf are better preserved in our scheme than in [4]. A well-recognized fact in lossy image coding is that longer filters gives better R-D performance shorter filters. The preference of longer filters can also be justified in the scenario of image denoising. As we can see the comparison results among 9/7, 10/18 and 28/28 filters from table 2, 28/28 filters achieves the smallest MSE among all.



Fig. 3. Comparison of reconstructed Barbara in the case of noise variance $\sigma = 25$. Top-left: original image. Top-right: noisy image. Bottom-left: SANS [4]. Bottom-right: Ours.

5. REFERENCES

- [1] D.L. Donoho and I.M. Johnstone, "Ideal spatial adaptation via wavelet shrinkage", *Biometrika*, Vol. 81, pp. 425-455, 1994
- [2] M.K. Mihcak, I. Kozintsev and K. Ramchandran, "Local statistical modeling of wavelet image coefficients and its application to denoising", *Proceeding of ICASSP 1999*, pp.3253-3256, March 1999
- [3] J. Liu and P. Moulin, "Image denoising based on scale-space mixture modeling of wavelet coefficients", *Proceeding of ICIP'99*, Kobe, Oct. 1999
- [4] S.G. Chang, B. Yu and M. Vetterli, "Spatially adaptive wavelet thresholding with context modeling for image denoising", *Proceeding of ICIP 1998*, Chicago, Oct. 1998
- [5] R.R. Coifman and D.L. Donoho, "Translation-Invariant denoising", *Wavelets and Statistics*, Springer-Verlag Lecture Notes, 1995
- [6] I.M. Johnstone and B.W. Silverman, "Wavelet threshold estimators for data with correlated noise", *Journal of Royal Statist. Soc. B* 59, pp.319-351, 1997
- [7] Xin Li and Michael T. Orchard, "Wavelet Image Coding via Rate-Distortion Optimized Backward Adaptive Estimation and Classification", submitted to *IEEE Trans. on Image Processing*, Jan. 2000