Wavelet Image Denoising Using Localized Thresholding Operators

M. Ghazel[1], G.H. Freeman[2], E.R. Vrscay[3], and R.K. Ward[1]

[1] Department of Electrical and Computer Engineering,
University of British Columbia, Vancouver, BC V6T 1Z4
[2] Department of Electrical and Computer Engineering
[3] Department of Applied Mathemarics,
University of Waterloo, Waterloo, ON N2L 3G1

Abstract. In this paper, a localized wavelet thresholding strategy which adopts context-based thresholding operators is proposed. Traditional wavelet thresholding methods, such as VisuShrink, LevelShrink and BayesShrink, apply the conventional hard and soft thresholding operators and only differ in the selection of the threshold. The conventional soft and hard thresholding operators are point operators in the sense that only the value of the processed wavelet coefficient is taken into consideration before thresholding it. In this work, it will be shown that the performance of some of the standard wavelet thresholding methods can be improved by applying a localized, context-based, thresholding strategy instead of the conventional thresholding operators.

1 Introduction

Over the past decade, various wavelet-based methods have been proposed for the purpose of image enhancement and restoration. Basic wavelet image restoration techniques are based on thresholding in the sense that each wavelet coefficient of the image is compared to a given threshold; if the coefficient is smaller than the threshold, then it is set to zero, otherwise it is kept or slightly reduced in magnitude. The intuition behind such an approach stems from the fact that the wavelet transform is efficient at energy compaction, thus small wavelet coefficients are more likely due to noise, and large coefficients are generally due to important image features, such as edges. Most of the efforts in the literature have concentrated on developing threshold selection criteria. Originally, Donoho and Johnstone proposed the use of a *universal* threshold applied uniformly throughout the entire wavelet decomposition tree [3,4]. Then the use of different thresholds for different subbands and levels of the wavelet tree was found to be more efficient [5]. Some methods of selecting thresholds that adapt to different spatial characteristics have recently been proposed and investigated [1]. It was found that such adaptivity in the threshold selection tends to improve the wavelet thresholding performance because it accounts for additional local statistics of the image, such as smooth or edge regions.

M. Kamel and A. Campilho (Eds.): ICIAR 2005, LNCS 3656, pp. 149–158, 2005.

Although adaptive wavelet thresholding methods attempt to employ thresholds that are adaptive to the local characteristics of the wavelet coefficients of the signal, they still apply the conventional hard and soft thresholding operators. These thresholding operators are point operators which are applied on each wavelet coefficients independently of its location or context. While it is generally assumed that the wavelet transform performs a significant degree of decorrelation between neighboring pixels, it is evident that some degree of dependence between neighboring wavelet coefficients remains. Intuitively, it seems more reasonable that when thresholding a wavelet coefficient, other neighboring coefficients should also be taken into consideration.

In this paper, localized, context-based hard and soft thresholding operators, which take into consideration the content of an immediate neighborhood when thresholding a wavelet coefficient, are proposed. It will be experimentally shown that the performance of three traditional wavelet thresholding methods, namely *VisuShrink, LevelShrink* and *BayesShrink*, may be significantly improved by using these new context-based thresholding operators instead of the conventional point operators.

This paper is organized as follows: Standard wavelet thresholding for image denoising is briefly discussed in section 2. The localized thresholding operators are then introduced in section 3. Section 4 includes the use of the cycle spinning algorithm for enhancing the denoised estimates. Experimental results and concluding remarks are given in sections 5 and 6, respectively.

2 Wavelet Thresholding for Image Denoising

Standard wavelet thresholding can be performed in the following steps:

1. Compute a linear forward discrete wavelet transform of the noisy signal.
2. Perform a nonlinear thresholding operation on the wavelet coefficients of the noisy signal.
3. Compute the linear inverse wavelet transform of the thresholded wavelet coefficients.

The second step in the above wavelet thresholding algorithm involves the selection of the threshold, λ, and the application of a thresholding operator. While the selection of the threshold differs from one method to another, most traditional wavelet thresholding methods apply the conventional hard and soft thresholding operators, described next.

2.1 Conventional Thresholding Operators

Traditional wavelet thresholding methods have adopted the following conventional operators corresponding to a threshold λ,

– The *hard thresholding* operator is defined as:

$$\hat{\mathbf{X}} = T_h(\mathbf{Y}, \lambda) \text{ such that } \hat{x}_{ij} = T_h(y_{ij}, \lambda) = \begin{cases} y_{ij} & \text{if } |y_{ij}| \geq \lambda, \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

– The *soft thresholding* operator is defined as:

$$\hat{\mathbf{X}} = T_s(\mathbf{Y}, \lambda) \text{ such that } \hat{x}_{ij} = T_s(y_{ij}, \lambda) = \begin{cases} y_{ij} - \lambda & \text{if } y_{ij} \geq \lambda, \\ y_{ij} + \lambda & \text{if } y_{ij} \leq -\lambda, \\ 0 & \text{otherwise,} \end{cases} \quad (2)$$

where $\mathbf{X} = [x_{ij}], \mathbf{Y} = [y_{ij}]$ and $\hat{\mathbf{X}} = [\hat{x}_{ij}]$ denote the wavelet coefficients of the original noiseless image, the noisy image and the denoised estimate, respectively.

The main difference among standard wavelet thresholding methods lies in the selection of the threshold λ. Next, three commonly known traditional wavelet thresholding methods are briefly described.

2.2 Standard Wavelet Thresholding Methods

In this section, three increasingly adaptive standard wavelet thresholding methods are reviewed.

- *VisuShrink:* Consists of applying the above thresholding operators using the universal threshold $\lambda_{univ} = \sqrt{2\ln(M)}\sigma_{\mathbf{w}}$, for a noisy signal of size M and noise intensity σ_w [3,4].
- *LevelShrink:* Account for some of the variability within the wavelet tree structure by using different thresholds for different decomposition levels. More specifically, a level-dependent threshold is given by $\lambda_j = \sqrt{2\ln(M)} \times \sigma_{\mathbf{w}} \times 2^{-(J-j)/2}$, for $j = 1, 2, \ldots J$, where J is the total number of decomposition levels and j is the scale level where the wavelet coefficient to be thresholded is located [5].
- *BayesShrink:* Assumes that the probability distribution of the noiseless wavelet coefficients follows a generalized Gaussian distribution [1]. An estimate of the optimal threshold, λ, is then selected by minimizing the mean-squared error between the noiseless wavelet coefficients, $\mathbf{X}$, and their denoised estimates, $\hat{\mathbf{X}}$. Based on this model for wavelet coefficients, it was experimentally shown that the following threshold:

$$\hat{\lambda}_j^{sub*} = \begin{cases} \frac{\sigma_{\mathbf{w}}^2}{\sqrt{\sigma^2_{\mathbf{Y}_j^{sub}} - \sigma^2_{\mathbf{w}}}} & \text{if } \hat{\sigma}^2_{\mathbf{Y}_j^{sub}} >> \hat{\sigma}^2_{\mathbf{w}}, \\ \max_{m=1,2,\ldots,M_j}\{|Y_{j,m}^{sub}|\} & \text{otherwise,} \end{cases} \quad (3)$$

is near optimal [1].

2.3 Remarks

The three standard wavelet thresholding methods described above apply the conventional hard and soft thresholding operators, defined in Eqs. (1) and (2). For a given threshold λ, these operators are point operators which are applied on each wavelet coefficient independently of their location or context. Consequently, the value of each thresholded wavelet coefficient depends only on the value of its noisy counterpart. While it is generally assumed that the wavelet transform performs a significant degree of decorrelation between neighboring pixels, it is well appreciated that some degree of dependence between neighboring wavelet coefficients remains. In fact, natural images' structures generally demonstrate similarities across a number of resolution scales of their wavelet coefficients. For instance, wavelet coefficients corresponding to a high activity subregion (i.e., edges) are generally clustered together and copied across the various resolutions and subbands of the wavelet tree. One should therefore expect some degree of dependence between neighboring wavelet coefficients corresponding to high activity subregions of the image. Thus, a more efficient thresholding operator should take advantage of this type of redundancy among neighboring wavelet coefficients.

3 Localized Wavelet Thresholding Operators

We now describe a context-based thresholding strategy that thresholds a noisy wavelet coefficient based not only on its value but also on the values of some of its neighbors. This method can be outlined as follows:

1. For each wavelet coefficient to be thresholded, y_{ij}, its neighborhood consists of an $m \times m$ mask centered at (but excluding) y_{ij} and is denoted by $\mathcal{C}_{m\times m}(y_{ij})$.
2. The maximum value (in magnitude) of the neighboring wavelet coefficients within the mask, $M_{ij} = \max_{\{(k,l)\neq(i,j)\}\in\mathcal{C}_{m\times m}(y_{ij})} |y_{kl}|$, is then computed.
3. Now for a given threshold λ, consider the following localized hard and soft thresholding operators:
   - The localized hard thresholding operator is defined by $\hat{\mathbf{X}} = T_h^{loc}(\mathbf{Y}, \lambda)$, such that:

$$\hat{x}_{ij} = T_h^{loc}(y_{ij}, \lambda) = \begin{cases} y_{ij} & \text{if } |y_{ij}| \geq \lambda \text{ OR } M_{ij} \geq \lambda, \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

   - The localized soft thresholding operator is defined by $\hat{\mathbf{X}} = T_s^{loc}(\mathbf{Y}, \lambda)$, such that:

$$\hat{x}_{ij} = T_s^{loc}(y_{ij}, \lambda) = \begin{cases} y_{ij} - \lambda & \text{if } y_{ij} \geq \lambda, \\ y_{ij} + \lambda & \text{if } y_{ij} \leq -\lambda, \\ y_{ij} & \text{if } |y_{ij}| < \lambda \text{ AND } M_{ij} \geq \lambda, \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

These simple, localized and context-based thresholding operators are presented as alternatives to the conventional hard and soft thresholding operators defined in Eqs. (1) and (2). The selection of these localized thresholding operators is explained and motivated in the following observations:

- These localized thresholding operators clearly take the values of the neighboring coefficients, located within the defined mask, into consideration before thresholding each wavelet coefficients. Thus, taking advantage of the dependence among neighboring wavelet coefficients.
- Note that only those wavelet coefficients that are insignificant and also surrounded by insignificant coefficients are set to zero. However, an insignificant coefficient is kept unchanged if it is located near a significant one.
- The issue of selecting the neighborhood and its size was investigated. It was observed that larger masks result in sharper, but noisier estimates, exhibiting more artifacts. On the other hand, smaller masks yield results that are closer to the standard wavelet thresholding methods. It was experimentally observed that a 3×3 window yields the best results for the three thresholding methods, studied in the previous section.

4 Enhancement Using Cycle Spinning

The denoised estimates obtained by wavelet thresholding methods often exhibit disturbing visual artifacts. In particular, pseudo-Gibbs phenomena tend to be noticeable in the vicinity of edges and other sharp discontinuities. The idea of using the cycle spinning algorithm has been previously proposed for the purpose of reducing the pseudo-Gibbs disturbing artifacts that are often present in wavelet-based image reconstruction and denoised estimates [2]. This procedure may be summarized as follows:

$$\hat{\mathbf{x}}_K = \frac{1}{K} \sum_{h=0}^{K-1} D_{-h}(IDWT(T_\lambda(DWT(D_h(\mathbf{y})))), \quad (6)$$

where the noisy image, $\mathbf{y}$, is first shifted, using a diagonal shifting operator, D_h, the DWT is then computed, and the thresholding method of choice, T_λ, is then applied. Then the IDWT is computed and the denoised image is unshifted. This process is repeated for each shift, $k = 1, 2, \ldots, K$, and the respective results are then averaged to obtain one denoised and enhanced estimate of the image.

5 Experimental Results

First, we present some results before and after the incorporation of the cycle spinning idea, for the commonly used test image of *Lena* (512×512 pixel, 8 bits per pixel) and its noisy observation as corrupted by an AWGN noise with intensity $\sigma_w = 25$, as illustrated in Fig. 1. We then present additional experimental results using various noisy images.

5.1 Before Cycle Spinning

Table 1 illustrates a quantitative comparison of the quality of the denoised estimates obtained by the three wavelet thresholding methods using the conventional

Original image
512 × 512 pixels

Noisy test image: $\sigma_w = 25$
RMSE=25.00, PSNR=20.17

Fig. 1. The original and the noisy version of the test image of *Lena*

Table 1. Conventional vs. context-based thresholding comparison

	Conventional thresholding				**Context-based thresholding**			
	Hard		Soft		Hard		Soft	
	RMSE	PSNR	RMSE	PSNR	RMSE	PSNR	RMSE	PSNR
VisuShrink	12.37	26.28	15.76	24.18	10.21	27.95	14.16	25.11
LevelShrink	10.01	28.11	11.30	27.07	*9.37*	*28.69*	10.07	28.07
BayesShrink	10.07	28.07	9.93	28.19	10.02	28.11	9.02	29.02

as well as the context-dependent hard and soft thresholding operators. In view of these results, we make the following observations:

- For the studied wavelet thresholding methods, there is an improvement in the quality of the denoised estimates obtained using the localized thresholding operators compared to the denoised images obtained by traditional thresholding schemes.
- The improvement achieved by the proposed localized thresholding operators is more evident for the case of the *VisuShrink* and *LevelShrink* than it is for the *BayesShrink* method. This can be explained by the fact that the optimal threshold corresponding to BayesShrink was originally derived specifically for the purpose of conventional soft thresholding as defined in Eq. (2). Thus this threshold may no longer be optimal when using the localized thresholding operators.

5.2 After Cycle Spinning

The cycle spinning algorithm was incorporated in order to enhance the denoised estimates obtained by the studied wavelet thresholding methods using the con-

Table 2. Conventional vs. context-based thresholding comparison after incorporating the cycle spinning idea with $K = 16$ shifts

	Conventional thresholding				**Context-based thresholding**			
	Hard		Soft		Hard		Soft	
	RMSE	PSNR	RMSE	PSNR	RMSE	PSNR	RMSE	PSNR
VisuShrink	10.27	27.90	14.94	24.64	8.87	29.18	12.76	26.01
LevelShrink	8.34	29.70	10.61	27.61	*8.06*	*30.00*	8.71	29.33
BayesShrink	8.64	29.40	8.66	29.38	8.71	29.33	8.36	29.69

ventional as well as the context-based hard and soft thresholding operators and the results are illustrated in Table 2. In view of these results we make the following observations:

- When comparing the results illustrated in Tables 1 and 2, it is evident that the quality of the denoised images is significantly improved by using the cycle spinning method.
- Generally, the use of the proposed localized thresholding operators yields better results than using the conventional thresholding operators before and after incorporating the cycle spinning idea.
- Experimentally, it was observed that the quality of the denoised estimate improves significantly after only a few shifts and then becomes stable and little or no further gains are achieved through additional shifts. In our case, a total of $K = 16$ diagonal shifts were used.
- Fig. 2 illustrates the results obtained by the *LevelShrink* method using the conventional as well as the context-based hard thresholding operators, before and after incorporating the cycle spinning algorithm. The context-based *LevelShrink* hard thresholding method yields the best results before and after the cycle spinning algorithm.
- Clearly the cycle spinning algorithm may be rather computationally expensive. Indeed, when incorporating this algorithm with K shifts for any denoising method, the computational complexity is multiplied by a factor of K.

5.3 Additional Experimental Results

Fig. 3 illustrates the results of denoising four different test images, *Lena*, *Boat*, *Peppers* and *San Francisco*, which were corrupted by an AWGN noise with varying intensity; $\sigma_{\mathbf{w}} = 10, 20, 30$ and 40, using the *BayesShrink* method before and after the incorporation of the cycle spinning (C.S.) algorithm. Again, two versions of the *BayesShrink* scheme were implemented: the conventional *BayesShrink* technique which adopts the conventional soft thresholding operator and a Context-Based (C-B) *BayesShrink* scheme which applies the proposed context-based soft thresholding operator.

(a) Standard hard *LevelShrink* RMSE=10.01, PSNR=28.11

(b) Context-based hard *LevelShrink* RMSE=9.37, PSNR=28.69

Before cycle spinning

(c) Standard hard *LevelShrink* RMSE=8.34, PSNR=29.70

(d) Context-based hard *LevelShrink* RMSE=8.06, PSNR=30.00

After cycle spinning

Fig. 2. A sample of the results

In view of these experimental results, we make the following observations:

- The results obtained by the C-B *BayesShrink* wavelet thresholding method are consistently better than those obtained by the conventional *BayesShrink* scheme. Indeed, this is the case for all test images and noise intensities.
- This improvement is even more evident before using the cycle spinning idea than after applying this enhancement method. This is probably because the use of the cycle spinning idea has benefited both methods by reducing most

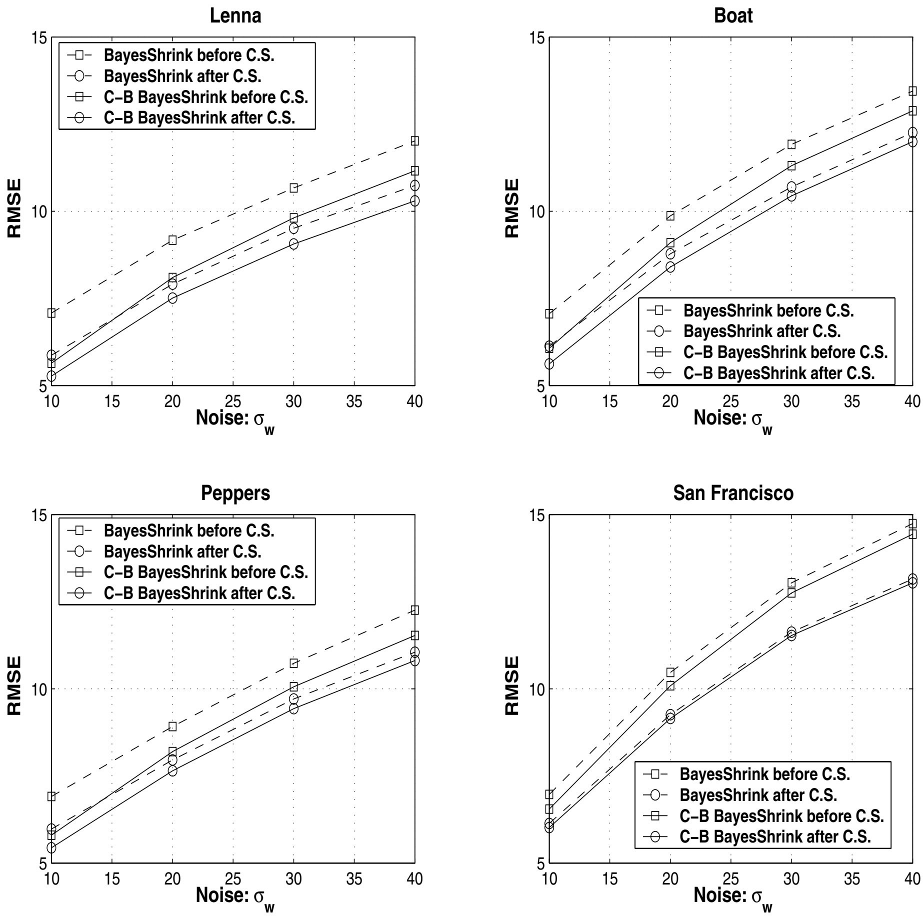

Fig. 3. Conventional vs. context-based (C-B) thresholding comparison for *BayesShrink* before and after applying the cycle spinning (C.S.) idea with $K = 16$ shifts

of the artifacts, hence yielding closer enhanced denoised estimates for both thresholding schemes.

– The cycle spinning idea is computationally expensive. In practice, some applications may not allow for this time complexity. Thus, the fact that the proposed context-based thresholding strategy yields significantly better results than the conventional thresholding operator without applying the cycle spinning idea is of great practical significance.

Based on the experimental results presented so far, it may be reasonable to conclude that, in general, the proposed context-based thresholding operators yield better results than the conventional thresholding operators.

6 Concluding Remarks

In this paper, we have proposed a generalized class of localized, context-based soft and hard wavelet thresholding operators that also depend upon the neighboring coefficient values. Our experiments have shown that these operators yield significant improvements over the conventional hard and soft thresholding point operators, especially for the *VisuShrink* and *LevelShrink* methods. The incorporation of cycle spinning further improves the results, but at significant computational expense.

References

1. S.G. Chang, B. Yu, and Martin Vetterli, "Spatially adaptive wavelet thresholding with context modeling for image denoising," *IEEE Trans. on Image Proc.*, vol. 9, no. 9, pp. 1522-1531, 2000.
2. R.R. Coifman, and D.L. Donoho, "Translation-invariant denoising," in A. Antoniadis, and G. Oppenheim, editors, Wavelets and Statistics, vol. 103 of Springer Lecture Notes in Statistics, pp. 125-150, New York, Springer-Verlag 1995.
3. D.L. Donoho, "Denoising and soft-thresholding," *IEEE Trans. Infor. Theory*, vol. 41, pp. 613-627, 1995.
4. D.L. Donoho, and I.M. Johnstone, "Ideal spatial adaptation via wavelet shrinkage," *Biometrika,* vol. 81, pp. 425-455, 1994.
5. S. Zhong, and V. Cherkassky, "Image denoising using wavelet thresholding and model selection," *Proc. IEEE Int. Conf. on Image Proc. (ICIP),* Vancouver, B.C., Sept. 2000.