

Review of Shrinkage Techniques for Image Denoising

Rajesh Kumar Rai
Associate Professor
ECE Department
NRIIST, Bhopal(M.P)

Jyoti Asnani
PG Scholar
ECE Department
NRIIST, Bhopal(M.P)

T. R. Sontakke
Principal
SCOE, Pune (M.S)

ABSTRACT

Image denoising is a common procedure to suppress the quality degradation caused by noise. Several image denoising methods are proposed in literature. Amongst these Discrete Wavelet Transform-DWT Filters are very popular. Denoising using the DWT-Transform includes decomposition of the image into various sub bands and then modeling them as independent identically distributed random variables with Gaussian distribution. Shrinkage methods are often used for suppressing Additive White Gaussian Noise (AWGN), where thresholding is used to retain the larger wavelet coefficients alone. Minimum Mean Square Error estimation is a common practice for noise analysis and is thus included in this paper. Overall we discuss in this review briefly the various Shrinkage methods in DWT-Domain Filters and we assess them by posing a comparison between the efficiency of these filters.

Keywords

PSNR, shrinkage methods, Thresholding, DWT-Domain Filters

1. INTRODUCTION

Image denoising algorithms are often used to enhance the quality of the images by suppressing the noise level while preserving the significant aspects of interest in the image. Several methods are proposed in literature for image denoising, where Discrete Wavelet Transform-Domain filters are very popular. A wavelet transform is the representation of a function by wavelets. The wavelets are scaled and translated copies (known as daughter wavelets) of a finite length or fast decaying oscillating waveform. A wavelet transform is classified into continuous wavelet transform (CWT) and DWT. The continuous wavelet transform (CWT) has received significant attention for its ability to perform a time-scale analysis of signals. On the other hand, the discrete wavelet transform (DWT) is an implementation of the wavelet transform using a discrete set of wavelet scales and translations obeying some definite rules. In other words, this transform decomposes the signals into mutually orthogonal set of wavelets. In this paper we only discuss methods based on DWT.

Recently, a lot of methods have been reported that perform denoising in DWT-domain. The transform coefficients within the sub bands of a DWT can be locally modeled as independent identically distributed (i.i.d.) random variables with generalized Gaussian distribution. Some of the denoising algorithms perform thresholding of the wavelet coefficients, which have been affected by additive white Gaussian noise, by retaining only large coefficients and setting the rest to zero. These

methods are popularly known as shrinkage methods. However, their performance is not quite effective as they are not spatially adaptive. Some other methods evaluate the de-noised coefficients by an Minimum Mean Square Error (MMSE) estimator, in terms of the noised coefficients and the variances of signal and noise. The signal variance is locally estimated by a Maximum Likelihood estimator (MLE) in small regions for every sub band where variance is assumed practically constant. These methods present effective results but they are not well suited near object edges where the variance field is not smoothly varied. Further, these methods introduce artifacts in the smooth regions of the output image. Some efficient wavelet-domain filters are discussed in subsequent subsections.

In this paper various shrinkage methods are reviewed. In section 2 various DWT-Domain based filters including shrinkage methods are defined. In section 3 a comparison is made between the aforementioned methods in section 2 based on existing research.

2. METHODS BASED ON DWT-DOMAIN FILTERS FOR IMAGE DENOISING

This section describes the various shrinkage methods used in literature for image de-noising in DWT-domain. The various shrinkage methods described in this section are based on estimation of statistical and probabilistic parameters for shrinkage.

2.1.1 VisuShrink

VisuShrink [1] is thresholding by applying universal threshold [2] proposed by Dohono and Johnston. This threshold is given by:

$$T_u = \sigma_n \sqrt{2 \log L}$$

Where, σ_n^2 is the noise variance of AWGN and L is the total number of pixels in an image. It is proved in [1] that a large fraction of any L number of random data array with zero mean and variance, σ_n^2 will be smaller than the universal threshold, T_u with high probability; the probability approaching 1 as L increases. Thus, with high probability, a pure noise signal is estimated as being identically zero. Therefore, for denoising applications, VisuShrink is found to yield a highly smoothed estimate. This is because the universal threshold is derived under the constraint that with high probability, the estimate should be at least as smooth as the signal. So the T_u tends to be high for large values of L, killing many signal coefficients along with the noise. Thus, the threshold does not adapt well to discontinuities in the signal.

2.1.2 SureShrink

SureShrink [3,4] is an adaptive thresholding method where the wavelet coefficients are treated in level-by-level fashion. In each level, when there is information that the wavelet representation of that level is not sparse, a threshold that minimizes Stein’s unbiased risk estimate (SURE) is applied. SureShrink is used for suppression of additive noise in wavelet-domain where a threshold T SURE is employed for denoising. The threshold parameter T SURE is expressed as:

$$T_{sure} = \arg \min_{T_h} (SURE(T_h; Y))$$

(SURE(T_h; Y)) is define by:

$$SURE(T_h; Y) = \sigma^2 - \frac{1}{L} \times \left[2\sigma_n^2, \left\{ i: |Y_i| \leq T_h - \sum_{i=1}^L \min(|Y_i|, T_h) \right\} \right]$$

Where,

σ_n^2 is the noise variance of AWGN;

L is the total number of coefficients in a particular sub-band;

Y_i is a wavelet coefficient in the particular sub-band.

2.1.3 BayesShrink

In BayesShrink [5], an adaptive data-driven threshold is used for image denoising. The wavelet coefficients in a sub-band of a natural image can be represented effectively by a Generalized Gaussian distribution (GGD). Thus, a threshold is derived in a Bayesian framework as:

$$T_B = \left(\frac{\hat{\sigma}_n^2}{\hat{\sigma}_F} \right)$$

where,

$\hat{\sigma}_n^2$ is the estimated noise variance of AWGN by robust median estimator and $\hat{\sigma}_F$ is the estimated signal standard deviation in wavelet-domain. The robust median estimator is stated as:

$$\sigma_n = \frac{\text{Median}[\{Y_{ij}\}]}{0.6745}, Y_{ij} \in \text{subband HH}$$

This estimator is used when there is no a priori knowledge about the noise variance.

2.1.4 OracleShrink and OracleThresh

OracleShrink and OracleThresh [5] are two wavelet thresholding methods used for image denoising. These methods are implemented with the assumption that the wavelet coefficients of original decomposed image are known. The OracleShrink and OracleThresh employ two different thresholds denoted as T_{OS} and T_{OT} respectively. Mathematically they are represented by:

$$T_{os} = \arg \min \sum_{i,j=1}^n \left(\xi_{T_h}(Y_{ij}) - F_{ij} \right)^2$$

$$T_{os} = \arg \min_{T_h} \sum_{i,j=1}^n \left(\zeta_{T_h}(Y_{ij}) - F_{ij} \right)^2$$

where, with {F_{ij}} are the wavelet coefficient of original decomposed image;

$\xi_{T_h}(\cdot)$ and $\zeta_{T_h}(\cdot)$ are soft-thresholding and hard-thresholding functions.

2.1.5 NeighShrink

Chen et al. proposed a wavelet-domain image thresholding scheme by incorporating neighboring coefficients, namely NeighShrink [6]. The method NeighShrink thresholds the wavelet coefficients according to the magnitude of the squared sum of all the wavelet coefficients, i.e., the local energy, within the neighborhood window. The neighborhood window size may be 3×3, 5×5, 7×7, 9×9, etc. But, the authors have already demonstrated through the results that the 3×3 window is the best among all window sizes. The shrinkage function for NeighShrink of any arbitrary 3×3 window centered at (i,j) is expressed as:

$$\Gamma_{ij} = \left(1 - \frac{T_U^2}{S_{ij}^2} \right)$$

Where, UT is the universal threshold and 2ij S is the squared sum of all wavelet coefficients in the respective 3×3 window given by:

$$S_{ij}^2 = \sum_{n=j-1}^{j+1} \sum_{m=i-1}^{i+1} Y_{m,n}^2$$

Here, + sign at the end of the formula means to keep the positive values while setting it to zero when it is negative. The estimated center wavelet coefficient F_{ij} is then calculated from its noisy counterpart Y_{ij} as:

$$\widehat{F}_{ij} = \Gamma_{ij} Y_{ij}$$

2.1.6 SmoothShrink

Mastriani et al. Proposed SmoothShrink [7], wavelet-domain image denoising method, for images corrupted with speckle noise. It employs a convolution kernel based on a directional smoothing (DS) function applied on the wavelet coefficients of the noisy decomposed image. The size of the window may vary from 3×3 to 33×33, but the studies [104] show that the 3×3 window gives better result as compared to others. Though this approach is meant for speckle noise, it is observed that it works satisfactorily even for additive noise. Therefore, the method: SmoothShrink is stated here.

2.1.6.1 SmoothShrink Algorithm

The SmoothShrink Algorithm is described in following steps:-

Step-1:

The average of the wavelet coefficients in four directions (d1, d2, d3, d4) as shown in Fig.1 is calculated.

Step-2:

The absolute difference between the center wavelet coefficient and each directional average is calculated as:

$$\Delta d_n = \left| \bar{Y}_{d_n} - Y_{ij} \right|, \quad n = 1, 2, 3, 4$$

Where, Y_{dn} = average of wavelet coefficients in nth direction.

Step-3:

The directional average which gives minimum absolute difference is found out.

$$k = \text{agr} \min_{\bar{Y}_{dn}} [\{|\Delta d_n|\}]$$

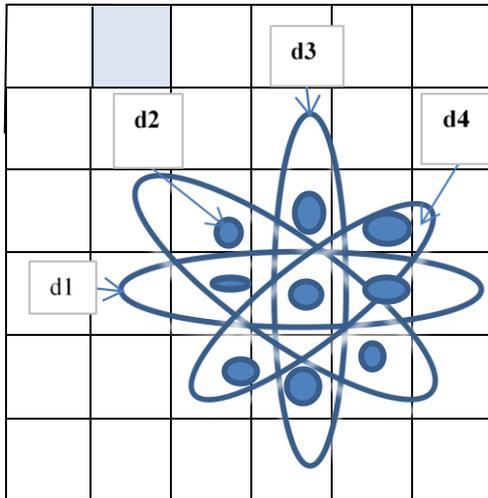


Fig. 1: 3x3 Directional smoothing window showing four directions d1, d2, d3, d4.

Step-4:

The estimated center wavelet coefficient is therefore replaced with the minimum directional average obtained in Step-3, i.e.:

$$\widehat{F}_{ij} = K$$

The SmoothShrink algorithm is applied to all sub bands of noisy decomposed image except the LL sub band.

2.1.6.2 BiShrink

Sendur et al. [9], [10]: The authors' method, called BiShrink, is based on new non-Gaussian bivariate distributions to model interscale dependencies. A nonlinear bivariate shrinkage function using the maximum *a posteriori* (MAP) estimator is then derived. In a second paper, these authors have extended their approach by taking into account the intrascale variability of wavelet coefficients. These techniques have been devised for both redundant and no redundant transforms.

2.1.6.3 ProbShrink

Pižurica et al. [8] Assuming a generalized Laplacian prior for the noise-free data, the authors' approach called ProbShrink is driven by the estimation of the probability that a given coefficient contains significant information- Notion of "signal of interest".



a) Lena b) Barbara



c) Goldhill

Fig. 2: The original test images with 512x512 pixels.



a) House b) Pepper

Fig. 3: The original test images with 256x256 pixels.

3. COMPARISON AND ASSESSMENT WITH EXISTING RESULTS

TABLE 1(a)

Image	Vishushrink	Sure shrink	Bayesshrink
Lena	PSNR	PSNR	PSNR
$\sigma=10$	30.56	33.4755	33.4106
$\sigma=20$	28.75	30.0724	30.2258
$\sigma=30$	26.78	28.3935	28.4901
$\sigma=35$	25.41	27.8293	27.8593
Barbara			
$\sigma=10$	25.72	30.6327	31.0322
$\sigma=20$	23.91	27.2961	27.2843
$\sigma=30$	22.61	25.0969	25.2842
$\sigma=35$	21.99	24.2202	24.5200
Goldhill			
$\sigma=10$	28.64	31.8715	31.9004
$\sigma=20$	26.93	28.4362	28.6570
$\sigma=30$	24.84	27.0256	27.1133
$\sigma=35$	23.70	26.3356	26.6088

The existing spatial-domain filters: wavelet-domain thresholding: VisuShrink, Sure Shrink, BayesShrink, OracleShrink, NeighShrink, SmoothShrink, are simulated on MATLAB 7.5 platform. The test images: Lena, Goldhill and Barbara of sizes 512x512 corrupted with AWGN of standard

deviation $\sigma_n = 10, 20, 30, 35$. While Bishrink and ProbShrink are simulated on MATLAB 7.5 platform. The test images Pepper and House of sizes 256×256 corrupted with AWGN of standard deviation $\sigma_n = 10, 20, 30$.

TABLE 1 (b)

Image	Oracle shrink & Oracle thresh	Neigh shrink	Smooth shrink
Lena	PSNR	PSNR	PSNR
$\sigma=10$	33.6114/32.6988	34.45	30.41
$\sigma=20$	30.3813/29.5232	30.11	27.43
$\sigma=30$	28.6009/27.7106	27.69	24.88
$\sigma=35$	27.9492/27.0696	26.76	23.8
Barbara			
$\sigma=10$	31.5070/30.4935	32.92	25.87
$\sigma=20$	27.4079/26.3428	28.57	23.19
$\sigma=30$	25.3289/24.0979	26.11	20.84
$\sigma=35$	24.584/23.3580	25.27	19.82
Goldhill			
$\sigma=10$	31.9734/30.7980	33.05	26.81
$\sigma=20$	28.7682/27.7837	29.02	23.72
$\sigma=30$	27.1687/26.3061	26.75	21.19
$\sigma=35$	26.6525/25.7732	25.95	20.13

TABLE 2

Images	Bishrink	Prob shrink
Pepper (256X256)	PSNR VALUE	PSNR VALUE
$\sigma=10$	37.18	37.45
$\sigma=20$	33.28	30.31
$\sigma=30$	29.80	28.33
House(256x256)		
$\sigma=10$	34.71	35.15
$\sigma=20$	31.63	32.19
$\sigma=30$	29.83	30.38

4. CONCLUSION

The peak-signal-to-noise ratio (PSNR) and execution time are taken as performance measures. The PSNR values of the different filters for various images are given in the tables: Table-I (a, b). The highest (best) PSNR value for a particular standard deviation of Gaussian noise is highlighted to show the best performance. The filtering performance is better if the method noise is very low since it talks of little distortion when a non-noisy image is passed through a filter. Therefore, a least value of method noise for a particular noise standard deviation is highlighted to show the best performance. The filter having less execution time is usually required for online and real-time applications. The least value of execution time is highlighted. The result implies the highest PSNR value is at lowest standard deviation and lowest PSNR value at highest standard deviation.

5. REFERENCES

- [1] D.L. Donoho, I.M. Johnstone, "Ideal spatial adaptation via wavelet shrinkage", *Biometrika*, vol. 81, no. 3, pp. 425- 55, 1994.
- [2] D.L. Donoho, I.M. Johnstone, "De-noising by soft-thresholding," *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 613-27, 1995.
- [3] D.L. Donoho, I.M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage", *Journal of the American Statistical Association*, vol. 90(432), pp. 1200-1224, 1995
- [4] F. Luisier, T. Blu, M. Unser, "A new SURE approach to image denoising: interscale orthonormal wavelet thresholding", *IEEE Transactions on Image Processing*, vol. 16 (3), pp. 593-606, 2007.
- [5] S. Chang, B. Yu, M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression", *IEEE Transactions on Image Processing*, vol. 9, no. 9, pp. 1532-1546, 2000.
- [6] G.Y. Chen, T.D. Bui, A Krzyzak, "Image denoising using neighboring wavelet coefficients", *proceedings of IEEE international conference on Acoustics, speech and signal processing '04*, pp. 917-920, 2004.
- [7] M Mastriani and A.E. Giraldez, "Smoothing of coefficients in wavelet domain for speckle reduction in synthetic aperture radar images", *Journal of Graphics Vision and Image Processing*, vol. 7(special issue), pp. 1-8, 2007.
- [8] A. Piżurica and W. Philips, "Estimating the probability of the presence of a signal of interest in multiresolution single and multiband image denoising," *IEEE Trans. Image Process.*, vol. 15, no. 3, pp. 654–665, Mar. 2006.
- [9] L. Sendur and I. W. Selesnick, "Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency," *IEEE Trans. Signal Processing*, vol. 50, no. 11, pp. 2744–2756, Nov. 2002.
- [10] L. Sendur and I. W. Selesnick, "Bivariate shrinkage with local variance estimation," *IEEE Signal Process. Lett.* vol. 9, no. 12, pp. 438–441, Dec. 2002.