Adaptive Wavelet Thresholding for Image Denoising Using Various Shrinkage Under Different Noise Conditions

Namrata Dewangan, Agam Das Goswami
Department of Electrical & Electronics, Shri Shankaracharya Group of Institutions, Bhilai

Abstract

This paper presents a comparative analysis of different image denoising thresholding techniques using wavelet transforms. There are different combinations that have been applied to find the best method for denoising. Visual information transmitted in the form of digital images is becoming a major method of communication, but the image obtained after transmission is often corrupted with noise. . The search for efficient image denoising methods is still a valid challenge at the crossing of functional analysis and statistics. Wavelet algorithms are useful tool for signal processing such as image compression and denoising. Image denoising involves the manipulation of the image data to produce a visually high quality image. The main aim is to modify the wavelet coefficients in the new basis, the noise can be removed from the data. In this paper, we analyzed several methods of noise removal from degraded images with Gaussian noise and Speckle noise by using adaptive wavelet threshold (Neigh Shrink, Sure Shrink, Bivariate Shrink and Block Shrink) and compare the results in term of PSNR and MSE.

Keywords— wavelet thresholding, Neigh Shrink, Sure Shrink, Bivariate Shrink and Block Shrink

1. Introduction

An image is corrupted by noise in its acquisition and transmission. The goal of image denoising is to produce good quality of the original image from noisy image. Wavelet denoising techniques remove the noise present in the signal while preserving the signal characteristics, regardless of its frequency content. De-noising of natural images corrupted by noise using wavelet techniques is very effective because of its ability to capture the energy of a signal in few energy transform values. Wavelet thresholding is a technique that exploits the capabilities of wavelet transform for signal denoising. It removes noise by killing coefficients that are insignificant relative to some threshold, and turns out to be simple and effective, depends on the choice of thresholding parameter and the choice of this threshold determines, to a great extent the efficacy of denoising. Simple de-noising algorithms that use the wavelet transform consist of three steps.

- Calculate the wavelet transform of the noisy signal.
- Modify the noisy wavelet coefficients according to some rule.
- Compute the inverse transform using the modified coefficients.

The problem of Image de-noising can be summarized as follows,

Let $A(i,j)$ be the noise-free image and $\mathrm{B}(i,j)$ the image corrupted with noise $Z(i,j)$

$$\mathrm{B}(i,j) = A(i,j) + SZ(i,j) \quad (1)$$

The problem is to estimate the desired signal as accurately as possible according to some criteria. In the wavelet domain, the problem can be formulated as

$$\mathrm{Y}(i,j) = W(i,j)\, Z(i,j) \quad (2)$$

Where $Y(i,j)$ is noisy wavelet coefficient; $\mathrm{W}(i,j)$ is true coefficient and $Z(i,j)$ noise.

The performance of the image de-noising algorithms has been investigated in terms of two parameters PSNR (peak signal to noise ratio) and MSE (mean square error).

2. Wavelet Thresholding

Let ***A*** = {A*ij*, i, j = 1, 2…M} denote the M×M matrix of the original image to be recovered, where M is some integer power of 2. During transmission the image is corrupted by independent, white Gaussian Noise Z*ij* with standard deviation _ i.e. *nij* ~ N (0, _2) and at the receiver end, the noisy observations B*ij*= *Aij* + Z*ij* is obtained. The goal is to estimate the signal ***A*** from noisy observations *Bij* such that Mean Squared error (MSE) is minimum. Let W and W^{-1} denote the two dimensional orthogonal discrete wavelet transform (DWT) matrix and its inverse respectively. Then Y = WB represents the matrix of wavelet coefficients of B having four sub bands (LL, LH, HL and HH) . The sub-bands HHk, HLk, LHk are called *details*, where k is the scale varying from 1, 2 …… J and J is the total number of decompositions. The size of the sub band at

scale k is N/2k × N/2k. The sub band LLJ is the lower resolution residue. The wavelet thresholding denoising method processes each coefficient of Y from the detail sub bands with a soft threshold function to obtain X . The denoised estimate is inverse transformed to A= W^{-1} X. In the experiments, soft thresholding has been used over hard thresholding because it gives more visually pleasant images as compared to hard thresholding; reason being the latter is discontinuous and yields
abrupt artifacts in the recovered images especially when the noise energy is significant.

2.1. Sure Shrink

Sure Shrink is a thresholding by applying sub-band adaptive threshold, a separate threshold is computed for each detail sub-band based upon SURE (Stein's unbiased estimator for risk), The goal of Sure Shrink is to minimize the mean squared error, defined as,

$$\text{MSE}=\frac{1}{n^2}\sum_{X,Y-1}^{n}(Z(X,Y)-S(X,Y))^2$$

(3)

Where Z(X,Y) is the estimate of the signal, S(X,Y) is the original signal without noise and n is the size of the signal. Sure Shrink suppresses noise by threshold the empirical wavelet coefficients. The Sure Shrink threshold t* is defined as

$$t^* = \min\,(t,\sigma\sqrt{2\log n})$$

(4)

Where t denotes the value that minimizes Stein's Unbiased Risk Estimator, σ is the noise variance and an estimate of the noise level σ was defined based on the median absolute deviation given by

$$\hat{\sigma} = \frac{median(\{|g_{j-1,k}|:k=0,1\ldots\ldots 2^{j-1}-1\})}{0.6745}$$

(5)

and n is the size of the image. It is smoothness adaptive, which means that if the unknown function contains abrupt changes or boundaries in the image, the reconstructed image also does.

2.2. Bivariate Shrinkage

Bivariate shrinkage function depends on both, coefficient and its parent yield improved results for wavelet based image denoising. Let w_2 represent the parent of w_1 (w_2 is the wavelet coefficient at the same position as w_1, but at the next coarser scale.) Then

$$y_1=w_1+n_1$$
$$y_2=w_2+n_2$$

(6)

Where y_1 and y_2 are noisy observations of w_1 and w_2 and n_1 and n_2 are noise samples. Then we can write

$$\mathbf{Y = w + n}$$

(7) $y= (y_1, y2)$

$w= (w_1, w2)$

$n= (n_1, n2)$

Standard MAP estimator for **w** given corrupted **y** is

$$\widehat{w}(y) = arg^{max}{}_{w} P_{w/y}(w/y)$$

(8)

This equation can be written as

$$\widehat{w}(y) = \arg max_w\ [P_{w/y}(w/y).P_w(w)]$$

(9)

$$\widehat{w}(y) = \arg max_w\ [P_N(y-W).P_w(w)]$$

(10)

According to bays rule allows estimation of coefficient can be found by probability densities of noise and prior density of wavelet coefficient. We assume noise is Gaussian then we can write noise as

$$P_n(n) = \frac{1}{2\pi{\sigma_n}^2} * exp\left(-{n_1}^2+{n_2}^2+2{\sigma_n}^2\right)$$

(11)

Joint of wavelet coefficients

$$P_W(W) = \frac{3}{2\pi\sigma^2} *\ exp\left(-\sqrt{3}\sqrt{W1^2}+W2^2\right)/\sigma$$

(12)

We know from equation (7)

$$\widehat{w}(y) = arg\ max_w\left[log\left(P_n(y-W)\right)+log\left(P_W(w)\right)\right]$$

(13)

Let us define $f(w) = log\left(P_W(W)\right)$

Then using equation 5.16 and 5.17

$$\widehat{w}(y) = \arg max_w\left[-\frac{(y_1-w_1)^2}{2{\sigma_n}^2}-\frac{(y_2-w_2)^2}{2{\sigma_n}^2}+f(w)\right] \quad (14)$$

This equation is equivalent to solving following equations

$$y_1 - \frac{\widehat{w_1}}{{\sigma_n}^2} + f_1(\widehat{w}) = 0$$

(15)

$$y_2 - \frac{\widehat{w_2}}{{\sigma_n}^2} + f_2(\widehat{w}) = 0$$

(16)

Where f1 and f2 represents the derivatives of $f(w)$ with respect to w_1 and w_2 respectively.

We know $f(w)$ can be written as

$$f(w) = log\left(P_w(w)\right)$$

$$= log\left(\frac{3}{2\pi\sigma^2} * \exp(-\sqrt{3}\sqrt{W1^2} + W2^2)/\sigma\right)$$

$$= log\left(\frac{3}{2\pi\sigma^2} - (-\sqrt{3}\sqrt{W1^2} + W2^2)/\sigma\right)$$

(17)

From this

$$f_1(w) = -\sqrt{3}\frac{w_1}{\sigma\sqrt{w_1^2}} + w_2^2$$

$$f_2(w) = -\sqrt{3}\frac{w_2}{\sigma\sqrt{w_1^2}} + w_2^2$$

(18)

From equations (15), (16), (17) and (18) MAP estimator can be written as

$$z_{\widehat{w_1}} = \frac{\left(\sqrt{y_1^2+y_2^2} - \sqrt{3}\frac{\sigma_n^2}{\sigma}\right) + y_1}{\sqrt{y_1^2+y_2^2}}$$

(19)

2.3. Neigh Shrink

Neigh Shrink thresholds the wavelet coefficients according to 3the 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. A 3× 3 neighboring window centered at the wavelet coefficient to be shrinked is shown in Fig 1.

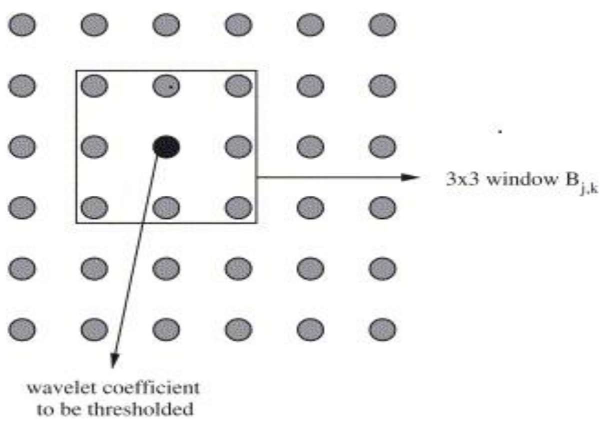

Fig.1. An illustration of the neighboring window of size 3× 3 centered at the wavelet coefficient to be shrinked.

The shrinkage function for Neigh Shrink of any arbitrary 3×3 window centered at (i,j) is expressed as:

$$\Gamma_{i,j} = [1 - \frac{T_U^2}{S_{ij}^2}]_+$$

(20)

Where T_U is the universal threshold and S_{ij}^2 is the squared sum of all wavelet coefficients in the given window.

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

(21)

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

$$\widehat{F_{ij}} = \Gamma_{ij} \cdot Y_{ij}$$

(23)

2.4. Block Shrink

Block Shrink is a data-driven block thresholding approach. It use the pertinence of the neighbor wavelet coefficients by using the block thresholding. It can decide the optimal block size and threshold for every wavelet subband by minimizing Stein's unbiased risk estimate (SURE). The block thresholding simultaneously keeps or kills all the coefficients in groups rather than individually. The block thresholding increases the estimation precision by utilizing the information about the neighbor wavelet coefficients. Unfortunately, the block size and threshold level play important roles in the performance of a block thresholding estimator. The local block thresholding methods mentioned above all have the fixed block size and threshold and same thresholding rule is applied to all resolution levels regardless of the distribution of the wavelet coefficients.

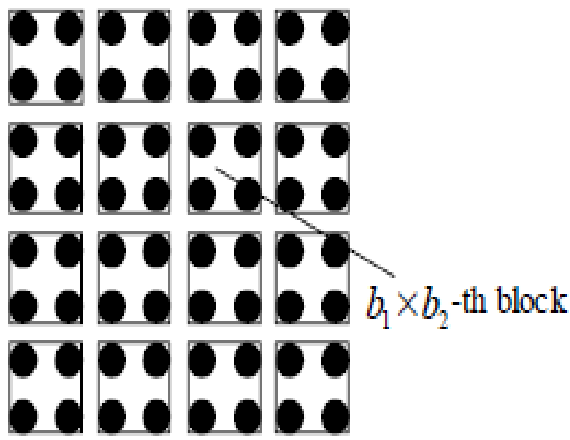

Figure 2: 2×2 Block partition for a Wavelet subband

As shown in Figure 2, there are a number of subbands produced when we perform wavelet decomposition on an image. For every subband, we need to divide it into a lot of square blocks.

Block Shrink can select the optimal block size and threshold for the given subband by minimizing Stein's unbiased risk estimate.
There are two parameters, PSNR (peak signal to noise ratio) and MSE (Mean Square Error) are calculated for all the standard images with their noisy and denoised images, respectively.
PSNR stands for the peak signal to noise ratio. It is used to calculate the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. Because many signals have a very wide dynamic range, PSNR is usually expressed in terms of the logarithmic decibel scale. It is most commonly used as a measure of quality of reconstruction in image compression etc. It is calculated as the following:

$$PSNR=10 \log (255/MSE)^{2} \qquad (24)$$

After the image is denoised, it is calculated for denoised image.
MSE indicates average error of the pixels throughout the image. The definition of a higher MSE does not indicate that the denoised image suffers more errors instead it refers to a greater difference between the original and denoised image. This means that there is a significant speckle reduction. The formula for the MSE calculation is given in equation

$$MSE = 1/N \sum_{j=0}^{N-1} (X_j - \bar{X}_j)^2 \qquad (25)$$

I. RESULT

This paper presents a comparative analysis of various image denoising techniques using wavelet transforms. The image formats that have been used in this work are JPG, BMP, TIF and PNG. We have experimented with four different thresholding methods (*Sure shrink, Bivariate shrink, Neigh shrink, Block Shrink)* using the various noisy images and report the results for the 512×512 standard test images *Lena* (Fig. 3). They are contaminated with Gaussian noise, salt and paper noise and speckle noise with standard deviations 10. Our results are measured by the PSNR and MSE.

Figure 3: 512×512 standard test images *Lena*

Table 1 Denoising results (PSNR) for *Lena* image

Thresholding Technique	Gaussian Noise	Salt & Paper Noise	Speckle Noise
Sure Shrink	27.07	18.42	19.62
Bivariate Shrink	75.66	74.52	75.69
Neigh Shrink	30.41	22.84	24.48
Block Shrink	67.54	63.54	58.23

Table 2 Denoising results (MSE) for *Lena* image.

Thresholding Technique	Gaussian Noise	Salt & Paper Noise	Speckle Noise

Sure Shrink	94.86	496.15	327.20
Bivariate Shrink	10.01	27.07	15.23
Neigh Shrink	59.13	325.97	231.85
Block Shrink	74.28	82.56	63.34

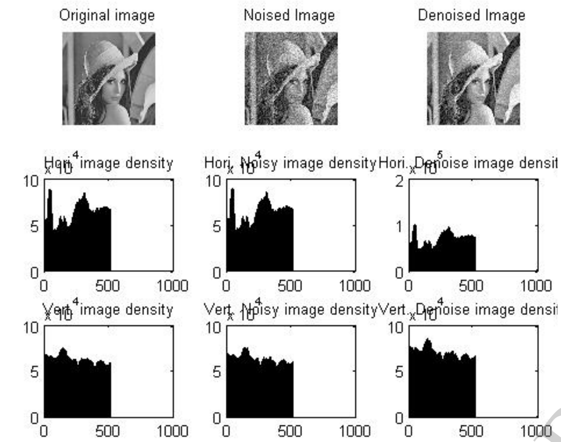

Fig.4. Matlab result of leena image with Gaussian noise using sure thresholding technique

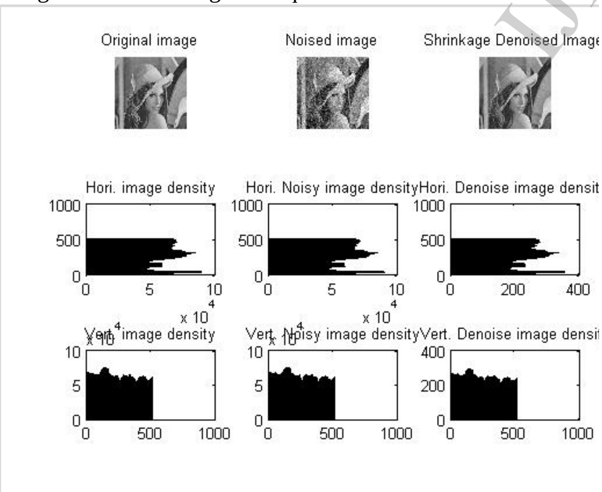

Fig.5. Matlab result of leena image with Gaussian noise using Bivariate thresholding technique

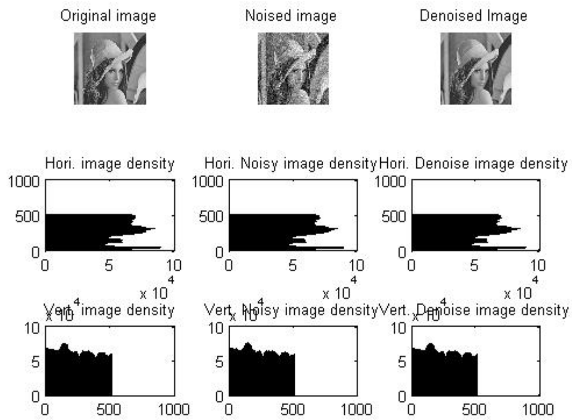

Fig.6. Matlab result of leena image with Gaussian noise using Neigh thresholding technique

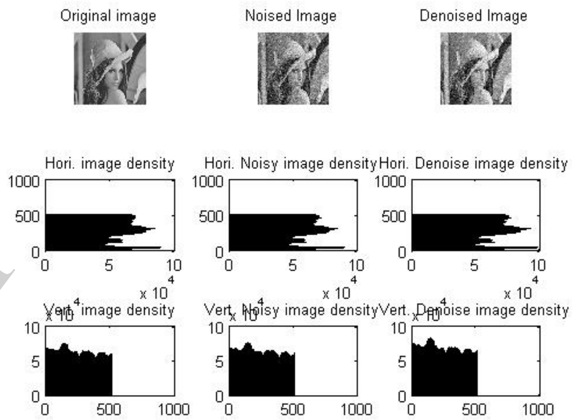

Fig.7. Matlab result of leena image with Gaussian noise using Block thresholding technique

REFERENCES

[1] Gao Zhing, Yu Xiaohai, "Theory and application of MATLAB Wavelet analysis tools", National defense industry publisher, Beijing, pp.108-116, 2004.

[2] Aglika Gyaourova Undecimated wavelet transforms for image denoising, November 19, 2002.

[3] C Sidney Burrus, Ramesh A Gopinath, and Haitao Guo, "Introduction to wavelet and wavelet transforms", Prentice Hall1997.S. Mallat, A Wavelet Tour of Signal Processing, Academic, New York, second edition, 1999.

[4] F.Abramovich and Y. Benjamini, "Adaptive thresholding of wavelet

coefficients,”*Comput. Statist. Data Anal.*, vol. 22, pp. 351–361, 1996.

[5] F. Abramovich, T. Sapatinas, and B. Silverman, “Wavelet thresholding via a Bayesian approach,” *J. R. Stat.*, vol. 60, pp. 725–749, 1998.

[6] Z. CAI, T. H. Cheng, C. Lu, and K. R. Subramanian, “Efficient waveletbased image denoising algorithm,” *Electron. Lett.*, vol. 37, no. 11, pp.683–685, May 2001.

[7] D. L. Donoho and I. M. Johnstone, ”Denoising by soft thresholding”,IEEE Trans. on Inform. Theory, Vol. 41, pp. 613-627, 1995.

[8] X.-P. Zhang and M. D. Desai, “Adaptive denoising based on SURE risk,” IEEE Signal Process. Lett., vol. 5, no. 10, pp. 265–267, Oct. 1998.

[9] Chang, S. G., Yu, B., and Vetterli, M. (2000). Adaptive wavelet thresholding for image denoising and compression. IEEE Trans. On Image Proc., 9, 1532–1546.

[10] G. Y. Chen, T. D. Bui And A. Krzyzak, Image Denoising Using Neighbouring wavelet Coefficients, Icassp ,Pp917-920

[11] D. L. Donoho and I. M. Johnstone, “Adapting to unknown smoothness via wavelet shrinkage,” *Journal of the American Statistical Assoc.*, 90(432), pp. 1200–1224, 1995.

[12] C. Stein, “Estimation of the mean of a multivariate normal distribution,” *Ann. Statist.*, 9, pp. 1135–1151, 1981.

[13] T. T. Cai, “Adaptive wavelet estimation, “A block thresholding and oracle inequality approach,” *Ann. Statist.*, 27, pp. 898-924, 1999.

[14] G. Y. Chen, T. D. Bui, A. Krzyzak, “Image denoising with neighbor dependency and customized wavelet and threshold,” *Pattern Recognition*, 38, pp. 115 – 124, 2005.

[15] T. T. Cai and H. H. Zhou, “A Data-Driven Block Thresholding Approach To Wavelet Estimation,” *Ann. Statist.*, accepted. http://www.stat.yale.edu/~hz68/.

[16] Alle Meije Wink and Jos B.T.M.Roerdink/ “Denoising functional MR Images – a comparison of wavelet denoising and gaussian smoothing” / *IEEE Transactions on Medical Imaging*, vol. 23, no. 3/ March 2004

[17] Byung-Jun Yoon and P. P. Vaidyanathan/ “Wavelet-based denoising by customized thresholding”

[18] Aleksandra Piˇzurica, Alle Meije Wink, Ewout Vansteenkiste, Wilfried Philips and Jos B.T.M. Roerdink/ “A review of wavelet denoising in MRI and ultrasound brain imaging”/ Submitted to *current medical imaging reviews*