IMAGE DENOISING USING WAVELET THRESHOLDING METHODS
Namrata Dewangan¹, Agam Das Goswami²
Department of Electrical & Electronics, Shri Shankaracharya Group of Institutions, Bhilai

Abstract
This paper presents a comparative analysis of various image denoising techniques using wavelet transforms. A lot of combinations have been applied in order to find the best method that can be followed for denoising intensity images. In this paper, we analyzed several methods of noise removal from degraded images with Gaussian noise by using adaptive wavelet threshold (Bayes Shrink, Neigh Shrink, Sure Shrink, Bivariate Shrink and Block Shrink) and compare the results in term of PSNR and MSE.

Keywords— wavelet thresholding, Bayes Shrink, Neigh Shrink, SureShrink, Bivariate Shrink and Block Shrink

Introduction
An image is often corrupted by noise in its acquisition and transmission. The goal of image denoising is to produce good estimates of the original image from noisy observations. Wavelet denoising attempts to remove the noise present in the signal while preserving the signal characteristics, regardless of its frequency content.

In the recent years there has been a fair amount of research on wavelet thresholding and threshold selection for signal de-noising [3], [5]-[6], [4], because wavelet provides an appropriate basis for separating noisy signal from the image signal. Wavelet thresholding is a signal estimation 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 heavily on the choice of a thresholding parameter and the choice of this threshold determines, to a great extent the efficacy of denoising.

* Corresponding Author
E. mail: manisha771@gmail.com

Wavelet Thresholding
Let $f = \{f_{ij}, \ i, j = 1, 2...M\}$ denote the MxM matrix of the original image to be recovered and M is some integer power of 2. During transmission the signal $f$ is corrupted by independent and identically distributed (i.i.d) zero mean, white Gaussian Noise $n$ with standard deviation $\sigma$ i.e. $n \sim N(0, \sigma^2)$ and at the receiver end, the noisy observations $g_{ij} = f_{ij} + \sigma n_{ij}$ is obtained. The goal is to estimate the signal $f$ from noisy observations $g_{ij}$ 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 $Y = Wg$ represents the matrix of wavelet coefficients of $g$ having four sub bands (LL, LH, HL and HH) [6], [7]. 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 $N/2^k \times N/2^k$. The sub band LLJ is the low-resolution residue. The wavelet thresholding denoising method processes each coefficient of $Y$ from the detail sub bands with a soft thresholding function to obtain $\hat{f}$. The denoised estimate is inverse transformed to $\hat{f} = W^{-1}Y$. 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.

A. Bayes Shrink (BS)
The Bayes Shrink method is effective for images including Gaussian noise. The observation model is expressed as follows:
Y = X + V

Here Y is the wavelet transform of the degraded image, X is the wavelet transform of the original image, and V denotes the wavelet transform of the noise components following the Gaussian distribution N(0, \sigma^2). Here, since X and V are mutually independent, the variances \sigma_Y^2, \sigma_X^2 and \sigma_V^2 of y, x and v are given by:

\[
\sigma_Y^2 = \sigma_X^2 + \sigma_V^2
\]

It has been shown that the noise variance \sigma_Y^2 can be estimated from the first decomposition level diagonal sub-band HH1 by the robust and accurate median estimator.

\[
\hat{\sigma}_Y^2 = \text{median}(|HH1|)^2
\]

The variance of the sub-band of degraded image can be estimated as:

\[
\hat{\sigma}_Y^2 = \text{median}(|HH1|)^2 \sum_{i,j} \hat{A}_{ij}
\]

Where \hat{A}_{ij} are the wavelet coefficients of sub-band under consideration, M is the total number of wavelet coefficient in that sub-band. The bayes shrink thresholding technique performs soft thresholding, with adaptive data driven, sub-band and level dependent near optimal threshold given by:

\[
T_{BS} = \begin{cases} 
\frac{\hat{\sigma}_Y^2}{\hat{\sigma}_Y^2} & \text{if } \hat{\sigma}_Y^2 < \hat{\sigma}_X^2 \\
\text{max}(|\hat{A}_{ij}|) & \text{otherwise}
\end{cases}
\]

\[
\hat{\sigma}_X = \sqrt{\text{max}(\hat{\sigma}_Y^2 - \hat{\sigma}_V^2, 0)}
\]

In the case, where \hat{\sigma}_X^2 \hat{\sigma}_V^2 \hat{\sigma}_Y^2 is taken to be zero, i.e. TBS \to \infty, or, in practice, T_{BS} = \text{max}(|\hat{A}_{ij}|), and all coefficients are set to zero.

**B. Neigh Shrink (NS)**

Let \( g = \{g_k\} \) will denote the matrix representation of the noisy signal. Then, with \( Wg \) denotes the matrix of wavelet coefficients of the signal under consideration. For every value of \( w_{ij} \), let \( B_{w} \) is a neighboring window around \( w_{ij} \), \( w_{ij} \) denotes the wavelet coefficient to be shrinked. The neighboring window size can be represented as \( L \), \( L \), where \( L \) is a positive odd number. A 3 \times 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 \times 3 centered at the wavelet coefficient to be shrinked.

Let

\[
S_{ij} = \sum_{(m,n) \in B_{w}} w_{mn}^2
\]

We omit the corresponding terms in the summation when the above summation has pixel indexes out of the wavelet sub-band range. The shrinked wavelet coefficient according to the neighshrink is given by this formula:

\[
w_{ij}' = w_{ij} \beta_{ij}
\]

The shrinkage factor \( \beta_{ij} \) can be defined as:

\[
\beta_{ij} = (1 - T_{UNI}^2 / S_{ij}^2)
\]

Here, the + sign at the end of the formula means to keep the positive value while set it to zero when it is negative and \( T_{UNI} \) is the universal threshold, which is defined as:

\[
T_{UNI} = \sqrt{2\sqrt{\text{ln}(n)}}
\]

Where \( n \) is the length of the signal.

Different wavelet coefficient sub-bands are shrinked independently, but the universal threshold
TUNI and neighboring window size L kept unchanged in all sub-bands. The estimated denoised signal \( f' = f^\prime_j \) is calculated by taking the inverse wavelet transform of the shrinked wavelet coefficients \( w^\prime_j \) i.e. \( f = W^\dagger w^\prime \).

### C. SureShrink

SureShrink 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), a method for estimating the loss \( \| \hat{y} - y \|^2 \) in an unbiased fashion. In our case let wavelet coefficients in the jth sub-band be \( \{ X_i : i = 1 \ldots d \} \), \( \hat{y} \) is the soft threshold estimator \( \hat{y} = \eta_j(X) \), we apply Stein’s result to get an unbiased estimate of the risk

\[
E[\| \hat{y} - y \|^2] = d \cdot 2 \sigma^2 \hat{\sigma}^2 + \sum_{i=1}^{d} \min \{ 2 \sigma_i^2, \frac{1}{n_i} \}
\]

SURE (\( t ; X \) = d - 2# \( i = 1 \ldots d \}) + \sum_{i=1}^{d} \min \{ 2 \sigma_i^2, \frac{1}{n_i} \}

For an observed vector \( x \) (in our problem, \( x \) is the set of noisy wavelet coefficients in a sub-band), we could find the threshold \( \hat{t}^\prime \) that minimizes SURE (\( t ; X \)) \( \hat{t}^\prime = \arg \min SURE(t ; X) \)

### D. Bivariate Shrink

New shrinkage function which depends on both coefficient and its parent yield improved results for wavelet based image denoising. Here, we modify the Bayesian estimation problem as to take into account the statistical dependency between a coefficient and its parent.

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

\[
y_1 = w_1 + n_1 \]
\[
y_2 = w_2 + n_2 \]

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

\[
y = w + n \]
\[
y = (y_1, y_2) \]
\[
w = (w_1, w_2) \]
\[
n = (n_1, n_2) \]

Standard MAP estimator for \( w \) given corrupted \( y \) is

\[
\hat{w}(y) = \arg \max_w P_w(w|y) \]

This equation can be written as

\[
\hat{w}(y) = \arg \max_w [P_w w^t H_w w] \]

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

\[
P_n(n) = \frac{1}{2\sqrt{\pi}} e^{-n^2/2} \]

Joint of wavelet coefficients

\[
P_w(w) = \frac{1}{2\sqrt{\pi}} e^{-w^2/2} \]

We know

\[
\hat{w}(y) = \arg \max_w [log(P_w(y - w))) + logP_n(w)] \]

Let us define \( f(w) = log(p_w(w)) \)

Then using equation 18 and 19

\[
\hat{w}(y) = \arg \max_w [\frac{(y - w_1)^2}{2\sigma^2} - \frac{(y - w_2)^2}{2\sigma^2} + f(w)] \]

This equation is equivalent to solving following equations

\[
y_1 - \frac{\sigma_1^2}{\sigma_2^2} + f_1(w_1) = 0 \]

\[
y_2 - \frac{\sigma_2^2}{\sigma_2^2} + f_2(w_2) = 0 \]

Here \( f_1 \) and \( f_2 \) represent the derivative of \( f(w) \) with respect to \( w_1 \) and \( w_2 \) respectively.

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

\[
f(w) = log(p_w(w)) = log(3/2\pi\sigma^2) + \frac{1}{2\sqrt{\pi}} e^{-w^2/2} \]

From this

\[
f_1(w_1) = -\sqrt{3}w_2/(\sigma_2^2 + w_2^2) \]
\[
f_1(w_1) = -\sqrt{3}w_2/(\sigma_2^2 + w_2^2) \]

\[
f_2(w_2) = -\sqrt{3}w_1/(\sigma_1^2 + w_1^2) \]
\[
f_2(w_2) = -\sqrt{3}w_1/(\sigma_1^2 + w_1^2) \]
From equations (20), (21), (22) and (23) MAP estimator can be written as

\[
\hat{\theta}_1 = \frac{\left(\sqrt{\hat{\theta}_1^2 + \hat{\theta}_2^2} - \sqrt{3} \frac{\hat{\theta}_1^2}{\sigma}\right) + \theta_2}{\sqrt{\hat{\theta}_1^2 + \hat{\theta}_2^2}}
\]

**E. Block Shrink**

Block Shrink is a completely data-driven block thresholding approach and is also easy to implement. It utilizes the pertinence of the neighbour wavelet coefficients by using the block thresholding scheme. 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, enjoys a number of advantages over the conventional term-by-term thresholding. The block thresholding increases the estimation precision by utilizing the information about the neighbor wavelet coefficients.

**PSNR**

PSNR stands for the peak signal to noise ratio. It is an engineering term 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. It is most commonly used as a measure of quality of reconstruction in image compression etc. It is calculated as the following:

\[
PSNR = 10 \log_2 (255/MSE)^2
\]

**MSE**

MSE indicates average error of the pixels throughout the image. In our work, a 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.

\[
MSE = \frac{1}{N} \sum_{j=0}^{N-1} (X_j - \hat{X}_j)^2
\]

**Result**

**Conclusion**

This thesis presents a comparative analysis of various image denoising techniques using wavelet transforms. A lot of combinations have been applied in order to find the best method that can be followed for denoising intensity images. The image
formats that have been used in this work are JPG, BMP, TIF and PNG.
We observe how wavelet transforms can be implemented to scale and translate a noise speckle image into a multi-resolution analysis representation. Bivariate shrinkage function used to reduce Gaussian, Salt & Pepper and speckle noise at different resolution levels. These results obtained have shown significant noise reduction then standard denoising methods such as Sure Shrink, ayes Shrink, Bivariate Shrinkage, Neigh Shrink and Block Shrink.

References
7. Savita Gupta and Lakhwinder kaur, Wavelet Based Image Compression using Daubechies Filters, In proc. 8th National conference on communications, I.I.T. Bombay, NCC-2002