Scaled Bayes Image Denoising Algorithm Using Modified Soft Thresholding Function

Sami M. A. Gorashi  
Computer Engineering Department  
Taif University-Khurmaha Branch  
KSA  

Saba Haldin A. Hussain  
Electrical and Computer Engineering Department  
Omdurman Islamic University  
Sudan  

Abstract

Image denoising is an active area of research and probably one of the most studied problems in the image processing fields. In this paper, a new adaptive threshold estimation and thresholding function are proposed for image denoising in wavelet domain. The proposed threshold estimation includes the effect of image subband bit size at each decomposition level in the calculation of the Bayesian based threshold value. The proposed thresholding function uses exponential weight to shrink the noisy coefficients. The experimental results show that the performance of the proposed denoising algorithm is superior to that of the conventional wavelet denoising approaches.

Keywords: Wavelet transform, Thresholding function, Image denoising.

1 Introduction

Digital images are generally affected by different types of noise. A noise is introduced in the transmission medium due to a noisy channel, imperfect instruments used in image processing, errors during the measurement process, noise due to degradation such as films, image compression and during quantization of the data for digital storage. Each element in the imaging chain such as lenses, film, digitizer, etc. contributes to the degradation. Thus, image denoising is a necessary and primary step in any further image processing tasks like segmentation, object recognition, computer vision, etc. To overcome image data corruption, we need to know something about the degradation process in order to develop a model for it. When we have a model for the degradation process, the inverse process can be applied to the image to restore it back to the original form. Noise modelling in images is greatly affected by capturing instruments, image quantization, data transmission channels,…etc. Mostly, noise in digital images is found to be additive in nature with uniform power in the whole bandwidth and with Gaussian probability distribution. Such a noise is referred to as Additive White Gaussian Noise (AWGN). White Gaussian noise can be caused by poor image acquisition or by transferring the image data in noisy communication channel. Most denoising algorithms use images artificially distorted with well-defined white Gaussian noise to achieve objective test results [4,14,15].

Denoising that based on wavelet transform for cancelling white Gaussian noise finds wide range of applications since the pioneer work by Donoho and Johnstone [7-9]. In wavelet based denoising algorithms, the noise is estimated and wavelet coefficients are thresholded to separate signal and noise using appropriate threshold value. Since the threshold plays a key role in this appealing technique, variant methods appeared later to set an appropriate threshold value [4,14,15]. Thresholding rules such as BayesShrink [4], and SUREShrink [7] are subband dependent, data-driven, and adaptive. In SUREShrink, a separate threshold is computed for each detail subband based upon minimizing Stein's Unbiased Risk Estimator (SURE) [16]. It is a hybrid of a universal and the SURE threshold, with the choice being dependent upon the energy of the particular subband. SUREShrink has yielded good image denoising performance and comes close to the true minimum MSE of the optimal soft-threshold estimator [7]. In BayesShrink, the assigned goal is the minimization of the Bayesian risk, and hence its name, BayesShrink. It minimizes the Bayes Risk Estimator function assuming Generalized Gaussian Distribution (GGD) prior. BayesShrink performs denoising that is consistent with the human visual system which is less sensitive to the presence of noise in vicinity of edges. It performs little denoising in high activity sub-region (image edges) to preserve the sharpness of the image but completely denoising the flat regions of the image [4]. In this paper, a Bayesian based denoising algorithm that uses scaled threshold according to the decomposed image subband bit size and modified soft thresholding function is proposed. Results show that the proposed denoising algorithm improves Peak Signal to Noise Ratio (PSNR) of the denoised image as compared with the conventional denoising algorithms.

The rest of the paper is organized as follows. Section 2 briefly reviews the basic idea behind discrete wavelet transform. Section 3 and section 4 explain different thresholding functions and different threshold estimation methods. In Section 5, we describe the proposed image denoising algorithm. The results of our proposed denoising algorithm will be compared with BayesShrink [4], VisuShrink [9], and spatial Wiener filter [10] in section 6. Finally, the concluding remarks are given in section 7.

2 Discrete Wavelet Transform

The wavelet transform was studied theoretically in geophysics and mathematics. Thereafter, Daubechies' [5], putting wavelets firmly into the application domain, pursued links with digital signal processing. Over the last two decades, the wavelet transform has acknowledged an
enormous deal of attention in many fields. In the fields of signal and image processing it is commonly used for denoising, compression, and image enhancement.

The main advantage of the discrete wavelet transform (DWT) is sparse representation. It is very helpful to describe local features, either spatially or spectrally over other image representations, such as the discrete Fourier transform. The sparse representation makes maximum noise removal while at the same time preserving the edges of the images or high frequency features of the signal which is very useful to analyze the image by experts. The DWT decomposes noisy image into four subbands namely LL, LH, HL, and HH. The decomposed noisy image consists of a small number of coefficients with high signal to noise ratio (SNR) named approximation subband (LL) and a large number of coefficients with low SNR named detail subbands (LH, HL, and HH). Thus, by thresholding the detail subbands, efficient noise suppression can be realized. Thresholding is a simple non-linear technique in which wavelet coefficient is thresholded by comparing against a threshold using appropriate thresholding function. Therefore, the thresholding function and the estimated threshold value play the major role in the denoising performance.

3 THRESHOLDING FUNCTION

3.1 HARD THRESHOLDING FUNCTION

To denoise a signal using Wavelet Transform(WT), the detail coefficients are thresholded. The simplest thresholding technique is the hard thresholding where the new values of the detail coefficients \( v \) are found according to the following thresholding function[8]:

\[
\theta_{\text{hard}}^T(v) = \begin{cases} 
  v & \text{if } |v| > T \\
  0 & \text{if } |v| \leq T
\end{cases}
\]  

(1)

Where \( T \) is the threshold value or the gate value. Notice that Eq.(1) is nonlinear and discontinuous at \( v = T \). The graphical representation of the hard threshold function is shown in figure (1).

![Figure 1. Hard Thresholding Function](image1.png)

The above figure clearly state that, all coefficients with magnitude greater than the selected threshold value \( T \) remain as they are and the others with magnitudes smaller than or equal to \( T \) are set to zero.

3.2 SOFT THRESHOLDING FUNCTION

Soft thresholding is another method of thresholding where the new values of the detail coefficients \( v \) are given by the following thresholding function[9]:

\[
\theta_{\text{soft}}^T(v) = \begin{cases} 
  \text{Sign}(v)(|v| - T) & \text{if } |v| > T \\
  0 & \text{if } |v| \leq T
\end{cases}
\]  

(2)

Generally, soft thresholding function is much better and yields more visually pleasant images compared with hard-thresholding[3,6]. This is because the hard-thresholding procedure creates discontinuities at \( T \) which yields abrupt artifacts in the recovered images. Also, the soft thresholding function yields a smaller minimum mean squared error compared to hard form of thresholding. The graphical representation of the soft thresholding is shown in figure(2).

![Figure 2. Soft Thresholding Function](image2.png)

3.3 MODIFIED SOFT THRESHOLDING FUNCTION

The thresholding function shape play significant role in any wavelet based denoising algorithm. Thus, several thresholding functions have been proposed by many authors all based on the conventional hard and soft thresholding function[2,11-13]. Both hard and soft thresholding functions, simply zeros all the coefficients which are smaller than the thresholding value. They assume that these coefficients are only due to noise which contradicts the fact that the desired signal is also present in these coefficients. Hence, this will cause loss some of image information. The degree of image information loss will depend upon the accuracy of the estimated threshold value that used to discriminate between noisy and noise free coefficients. To overcome this problem, we suggest to use modified soft thresholding function that uses exponential weight to shrink the coefficients which are smaller than the estimated threshold value \( T \) while keeping same scaling function that used for soft thresholding for the rest of coefficients according to the following proposed equation:

\[
\theta_{\text{modifiedsoft}}^T(v) = \begin{cases} 
  \text{Sign}(v)(|v| - T) & \text{if } |v| \geq T \\
  v \times e^{\delta(|v| - 3T/2)} & \text{if } |v| < T
\end{cases}
\]  

(3)

Where \( \delta \) is the parameter that controls the degree of wavelet coefficients falling down. Extensive simulation show that, for low noise level, \( \delta=0.09 \) is sufficient, while for high noise level, larger value of \( \delta (\delta \cong 0.8) \) is required to cancel heavy noise. Examining Eq.(3), clearly, as \( \delta \to \infty \), the thresholding rule follows soft thresholding function. Figure(3), shows the effect of \( \delta \) on the characteristics of the suggested modified soft thresholding function.
4 THRESHOLD ESTIMATION METHODS

4.1 VISUSHRINK

VisuShrink was introduced by Donoho[8,9]. It is non-adaptive universal threshold, which depends only on number of data points and the estimated noise standard deviation, and hence makes it very simple to implement. It follows the hard thresholding rule and also referred to as universal threshold which can be defined as:

\[ T_{\text{VisuShrink}} = \hat{\sigma}_n \sqrt{2\log(M)} \]  
\[ (4) \]

Where M represents the image size or number of pixels and \( \hat{\sigma}_n \) is an estimate of the noise level which can be calculated using Median Absolute Deviation (MAD) given by [8,9]:-

\[ \hat{\sigma}_n = \frac{\text{median}(|y_{ij}|)}{0.6745}, Y_{ij} \in HH_1 \]  
\[ (5) \]

Where \( Y_{ij} \) is the detail coefficients in the diagonal subband of the first decomposition level \( HH_1 \).

VisuShrink does not deal with minimizing the mean squared error [8]. It can be viewed as a general-purpose threshold selector that exhibits near optimal minimax error properties and ensures with high probability that the estimates are as smooth as the true underlying functions. However, the disadvantage of VisuShrink is that it yields recovered images that are overly smoothed. Because firstly, its threshold choice can be unwarrantedly large due to its dependence on the number of pixels in the image which in turn yields to remove too many wavelet coefficients of the decomposed image. Secondly, VisuShrink follows the global thresholding scheme where there is a single value of threshold applied globally to all wavelet coefficients.

4.2 BAYES SHRINK

Bayesian based threshold estimation was proposed by Chang, et al [4]. The goal of this method is to estimate a threshold value that minimizes the Bayesian risk assuming Generalized Gaussian Distribution (GGD) prior. It has been shown that BayesShrink outperforms SUREShrink most of the times in terms of PSNR values over a wide range of noisy images[4]. It uses soft thresholding and is subband-dependent, which means that thresholding is done at each band of resolution in the wavelet decomposition.

To estimate the threshold value, the corrupted image pixel can be written as (assume additive white Gaussian noise):

\[ r(x,y) = s(x,y) + n(x,y) \]  
\[ (6) \]

where \( r(x,y), s(x,y), \) and \( n(x,y) \) are the observed, original, and noise signals at spatial location \( (x,y) \) respectively.

Since the noise and the signal are independent of each other, it can be stated that:

\[ \sigma_r^2 = \sigma_s^2 + \sigma_n^2 \]  
\[ (7) \]

Where \( \sigma_r^2 \) is the observed signal variance, \( \sigma_s^2 \) is an estimate of the original noise free signal variance, and \( \sigma_n^2 \) is an estimate of noise variance. The noise standard deviation \( \sigma_n \) can be estimated from Eq.(5) while the observed signal variance \( \sigma_r^2 \) can be estimated using:

\[ \hat{\sigma}^2_r = \frac{1}{M^2} \sum_{x=1}^{M} r^2(x, y) \]  
\[ (8) \]

Knowing \( \hat{\sigma}_r^2 \) and \( \hat{\sigma}_n^2 \), the variance of the signal, \( \sigma_s^2 \) can be estimated according to:

\[ \hat{\sigma}_s^2 = \max(\hat{\sigma}_r^2 - \hat{\sigma}_n^2, 0) \]  
\[ (9) \]

Knowing \( \hat{\sigma}_s^2 \) and \( \hat{\sigma}_n^2 \), the Bayes threshold is estimated according to[4]:-

\[ T_{\text{Bayes}} = \frac{\hat{\sigma}_s^2}{\hat{\sigma}_n} \]  
\[ (10) \]

4.3 SCALED BAYES SHRINK

The Bayes Shrink threshold estimation can be improved by including the effect of image subband bit size at each decomposition level in the calculation of the threshold value. Empirically, based on extensive simulation test, we suggest to use the following formula for the scaled Bayes threshold estimation:

\[ T_{\text{ScaledBayes}} = (1 + \gamma) \frac{\hat{\sigma}_s^2}{\hat{\sigma}_n} \]  
\[ (11) \]

Where, \( \gamma = \frac{\text{Subband Bit Size}}{\text{Subband Bit Size}} \times \frac{2}{2} \) and \( \eta \) is the subband height.

The scaling parameter \( \gamma \) is a monotonically increasing value according to the level of decomposition. In general, most of wavelet coefficients are noisy at higher decomposition levels because the noise almost concentrated in high frequency bands. Accordingly, larger threshold value is necessary at higher decomposition level to cancel noise. This characteristic is met by the proposed \( T_{\text{ScaledBayes}} \) and hence improved denoising results is expected.

5 PROPOSED ALGORITHM

In the wavelet transform domain, several denoising algorithms were successfully applied in a wide range of applications. These algorithms were based on thresholding the wavelet detail coefficients. The selected threshold value and the thresholding function shape play significant role in deciding whether the coefficient is noisy(to cancel) or noise free(to keep or shrunk). In this paper, an improved Bayesian based threshold value is estimated according to Eq.(11) and a modified soft thresholding function according to Eq.(3) is used. The following steps describe the implementation of the proposed algorithm.

1. Assign the number of wavelet decomposition level.
2. Decompose the noisy image into four subbands namely LL, LH, HL, and HH using 2-D discrete wavelet transform.
3. Estimate the noise standard deviation $\sigma_n$ using Eq.(5).
4. Estimate the variance of the observed signal $\sigma^2$ in subbands LH, HL, and HH for each decomposition level using Eq.(8).
5. Estimate the variance of the desired signal $\sigma^2$ in subbands LH, HL, and HH for each decomposition level using Eq.(9).
6. For each decomposition level, calculate the scaling parameter $\gamma$.
7. Estimate scaled Bayes threshold $T_{\text{ScaledBayes}}$ in subbands LH, HL, and HH for each decomposition level using Eq.(11).
8. Denoise wavelet coefficients in subbands LH, HL, and HH for each decomposition level using modified soft thresholding function defined in Eq.(3).
9. Reconstruct the denoised image using inverse 2-D discrete wavelet transform.

As an illustration, Figure(4) shows a block diagram of the proposed denoising algorithm.

![Figure 4. Block Diagram of the Proposed Algorithm](image)

Where $W$ and $W^{-1}$ are the wavelet transform and its inverse respectively. $\hat{s}(x,y)$ is the recovered image pixel.

### 6 EXPERIMENTAL RESULTS

For evaluation purpose, an experiment was conducted to assess the performance of the proposed denoising algorithm. We will compare the performance of proposed algorithm with the spatial domain Wiener filter[10], VisuShrink[9], and BayesShrink[4] denoising algorithms. Images artificially corrupted with white Gaussian noise with zero mean and standard deviations 10, 15, 20, 25, and 30 were used in the test. The wavelet transform that employs Daubeche’s[5] least asymmetric compactly supported wavelet with eight vanishing moments was used for wavelet domain based denoising algorithms. The relative denoising algorithms performance is measured quantitatively using Peak Signal to Noise Ratio(PSNR) [1]. The denoising relative performance in terms of PSNR value for a set of images is recorded in Table(1). The best denoising algorithm is highlighted in bold font for each test image.

Subjectively, for low noise level degradation (see figure(5)), almost all denoising algorithms achieve nearly equivalent visual quality with the exception of VisuShrink which produces lower image quality. For higher noise density level, VisuShrink exhibits worst denoising results. This is due to the fact that, in the wavelet domain, the magnitude of the coefficients varies depending on the decomposition levels. Hence, if all levels are processed with just one threshold value the processed image may be overly smoothed so that sufficient information preservation is not possible and the image gets blurry. This is the nature of the VisuShrink denoising algorithm which uses single universal threshold value $T_{\text{Universal}}$ for all subbands in each level and hence produces an image of much lower resolution (as an evident see figure(6c), and figure(7c)). On the other hand, the proposed denoising algorithm removes noise and preserves image fine details better than both BayesShrink and Wiener filter. These algorithms are the nearest competitor to our proposed algorithm. In fact, the proposed algorithm uses more accurate estimate of threshold. Referring to Eq.(11), the proposed threshold estimate is monotonically increased in accordance with the level of decomposition. This will provide better noise cancellation of different frequency noise components in different decomposition levels. To demonstrate the effectiveness of the proposed algorithm in comparison with BayesShrink, and Wiener see figure(6,d,e,f) and figure(7,d,e,f) through noticing the sky, the face of Lena, and the edges of the images. Clearly, these figures state that, the proposed algorithm reveals better visual perception.

Table(1): PSNR Results for Denoising Lena, Mandrill, and Boat.

<table>
<thead>
<tr>
<th>Image</th>
<th>$\sigma_n$</th>
<th>VisuShrink</th>
<th>Weiner</th>
<th>BayesShrink</th>
<th>Proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lena</td>
<td>10</td>
<td>27.78</td>
<td>33.52</td>
<td>33.48</td>
<td>33.57</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>26.41</td>
<td>31.11</td>
<td>31.58</td>
<td>31.74</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>25.63</td>
<td>29.01</td>
<td>30.34</td>
<td>30.43</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>24.97</td>
<td>27.30</td>
<td>29.32</td>
<td>29.45</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>24.54</td>
<td>25.81</td>
<td>28.60</td>
<td>28.70</td>
</tr>
<tr>
<td>Mandrill</td>
<td>10</td>
<td>21.10</td>
<td>26.77</td>
<td>28.85</td>
<td>29.16</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>20.63</td>
<td>25.93</td>
<td>26.68</td>
<td>26.82</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>20.34</td>
<td>24.99</td>
<td>25.17</td>
<td>25.32</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>20.13</td>
<td>24.12</td>
<td>24.09</td>
<td>24.21</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>19.99</td>
<td>23.27</td>
<td>23.26</td>
<td>23.36</td>
</tr>
<tr>
<td>Boat</td>
<td>10</td>
<td>25.47</td>
<td>31.71</td>
<td>31.90</td>
<td>31.93</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>24.28</td>
<td>29.91</td>
<td>29.86</td>
<td>29.94</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>23.60</td>
<td>28.23</td>
<td>28.47</td>
<td>28.58</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>23.12</td>
<td>26.69</td>
<td>27.40</td>
<td>27.51</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>22.66</td>
<td>25.38</td>
<td>26.59</td>
<td>26.70</td>
</tr>
<tr>
<td>Average PSNR</td>
<td>23.52</td>
<td>27.68</td>
<td>28.51</td>
<td>28.62</td>
<td></td>
</tr>
</tbody>
</table>
7 CONCLUSIONS

In this paper, we can conclude that the proposed denoising algorithm outperforms other denoising algorithms objectively and subjectively. As noise level increased, VisuShrink exhibits worse denoising performance as compared with other denoising algorithms. It blurs the images while it attempts to remove noise. The proposed algorithm improves the PSNR of the image as compared with the Wiener and BayesShrink denoising algorithms. Referring to Table(1), the proposed algorithm achieves an average PSNR gain of 4.91 dB, 1.43 dB, 0.12 dB for Lena image; 5.09 dB, 1.1 dB, 0.11 dB for Mandrill image; and 5.1 dB, 0.94 dB, 0.11 dB for Boat image as compared with VisuShrink, Wiener, and BayesShrink respectively. The proposed algorithm can be extended to deal with three dimensional colour pictures through vector processing. This issue is left as future work.

References


