Comparative Study of Image Denoising Using Wavelet Transforms and Optimal Threshold and Neighbouring Window

Iman M.G. Alwan
ainms_66@yahoo.com
College of Education for Women - Computer Science Dept.

Abstract
NeighShrink is an efficient image denoising algorithm based on the discrete wavelet transform (DWT). Its disadvantage is to use a suboptimal universal threshold and identical neighbouring window size in all wavelet subbands. Dengwen and Wengang proposed an improved method, which can determine an optimal threshold and neighbouring window size for every subband by the Stein's unbiased risk estimate (SURE). Its denoising performance is considerably superior to NeighShrink and also outperforms SURE-LET, which is an up-to-date denoising algorithm based on the SURE. In this paper different wavelet transform families are used with this improved method, the results show that Haar wavelet has the lowest performance among other wavelet functions. The system was implemented using MATLAB R2010a. The average improvement in term of PSNR between Haar and other wavelet functions is 1.37dB.

Keywords: image denoising, wavelet transform, optimal threshold, neighbouring window, NieghShrink .

دراسة مقارنة لازالة الضوضاء من الصور باستعمال تحويلات المويجة و قيمة العتبة والنافذة المجاورة المثلى

ايمان محمد جعفر علوان
كلية التربية للبنات - قسم الحاسبات

الخلاصة
تعتبر خوارزمية (NeighShrink) لازالة الضوضاء خوارزمية كفوءة والتي تعتمد على تحويلة المويجة (DWT). ولكن من عيوب هذه الخوارزمية إعتمادها على قيمة عتبة (threshold) و حجم نافذة موحد لكل تقسيمات تحويلة المويجة. وقد تم إقتراح طريقة من قبل Dengwen و Wengang يمكن من خلالها تحديد قيمة العتبة وحجم النافذة تبعا لكل حزمة ترددات من تحويلة المويجة بالاعتماد على Stein's unbiased risk estimate (SURE). قدمت هذه الطريقة نتائج أفضل في إزالة الضوضاء. في هذا البحث تم تطبيق انواع مختلفة من تحويلة المويجة على هذا النظام ، وقد أظهرت النتائج إن أوطأ كفاءة هي باستعمال تحويلة (Haar) حيث كان الفرق في معدل التحسين بين هذه التحويلة و الانواع الاخرى 1.73 dB.

1- Introduction:

An image is often corrupted by noise in its acquition and transmission. Image denoising is used to remove the additive noise while retaining as much as possible the important signal features. In the recent years there has been a fair amount of research on wavelet thresholding and threshold selection for signal de-noising [1]. Both VisuShrink and SureShrink are the best known methods of wavelet shrinkage proposed by Donoho and Johnstone [2,3]. VisuShrink is simple and efficient. Its denoising procedures can be stated as follows. First perform a wavelet transform on the observed data which are corrupted by additive white Gaussian noise (AWGN), then apply soft or hard threshold to the wavelet detail coefficients using universal threshold, last obtain the denoised signal by performing the inverse wavelet transform to the thresholded wavelet coefficients. SureShrink denoised images include more significant wavelet coefficients thus alleviating the blurring problem produced by VisuShrink generating more detailed images. Many alternative methods have come forth. Cai and silverman proposed two different shrinkage methods NeighBlock and NeighCoeff for 1-D signals [4]. They threshold the wavelet coefficients in overlapping blocks rather than individually or term by term as VisuShrink or SureShrink. Chen et al. applied NeighCoeff to image denoising and their method is called NeighShrink [5]. Sender and Selesnick [6] use bivariate shrinkage function which models the statistical dependence between a wavelet coefficients and its parent. Luisier et al., [7] proposed a SURE-LET method which directly parametrize the denoising process as a sum of elementary nonlinear processes with unknown weights. It need not hypothesize a statistical model for the noiseless image while it minimizes an estimate of the mean squared error between the noiseless image and the denoised one by the SURE. Dengwen and Wengang [8] proposed a method which can estimate an optimal threshold and neighbouring window size for NeighShrink in every wavelet subband. In this paper, different wavelet types are applied for the algorithm in [8], to find the best one in terms of PSNR.

2- Wavelet Transform:

The wavelet expansion set is not unique. A wavelet system is a set of building blocks to construct or represents a signal or function. It is a two dimensional expansion set, usually a basis, for some class one or higher dimensional signals. The wavelet expansion gives a time frequency localization of the signal. Wavelet systems are generated from single scaling function by scaling and translation [9].

For the DWT special families of wavelet functions are developed. These wavelets are compactly supported, orthogonal or biorthogonal and are characterized by low-pass and high-pass analysis and synthesis filters. From the filters a wavelet function $\psi(t)$ and scaling function $\varphi(t)$ can be derived. Some generally used families for the DWT are The *Daubechies* familie is named after Ingrid Daubechies who invented the compactly supported orthonormal wavelet, making wavelet analysis in discrete time possible. The first order Daubechies wavelet is also known as the Haar wavelet, which wavelet function resembles a step function. The wavelet and scaling functions for the Daubechies functions with order 1 up to 8 are shown in respectively Fig.1-1 up to Fig.1-8. The Haar wavelet or db1 can be written as:

$$\psi(t) = \begin{cases} 1 & \text{if } x \in [0 \quad 0.5] \\ -1 & \text{if } x \in [0.5 \quad 1] \\ 0 & \text{if } x \notin [0 \quad 1] \end{cases}$$
$$\phi(t) = \begin{cases} 1 & \text{if } x \in [0 \quad 1] \\ 0 & \text{if } x \notin [0 \quad 1]. \end{cases} \qquad (1)$$

Higher order Daubechies functions are not easy to describe with an analytical expression. The order of the Daubechies functions denotes the number of vanishing moments, or the number of zero moments of the wavelet function. This is weakly related to the number of oscillations of the wavelet function. The larger the number of vanishing moments, the better the frequency localization of the decomposition. The dependence between wavelet coefficients on different scales decays with increasing wavelet order [10].

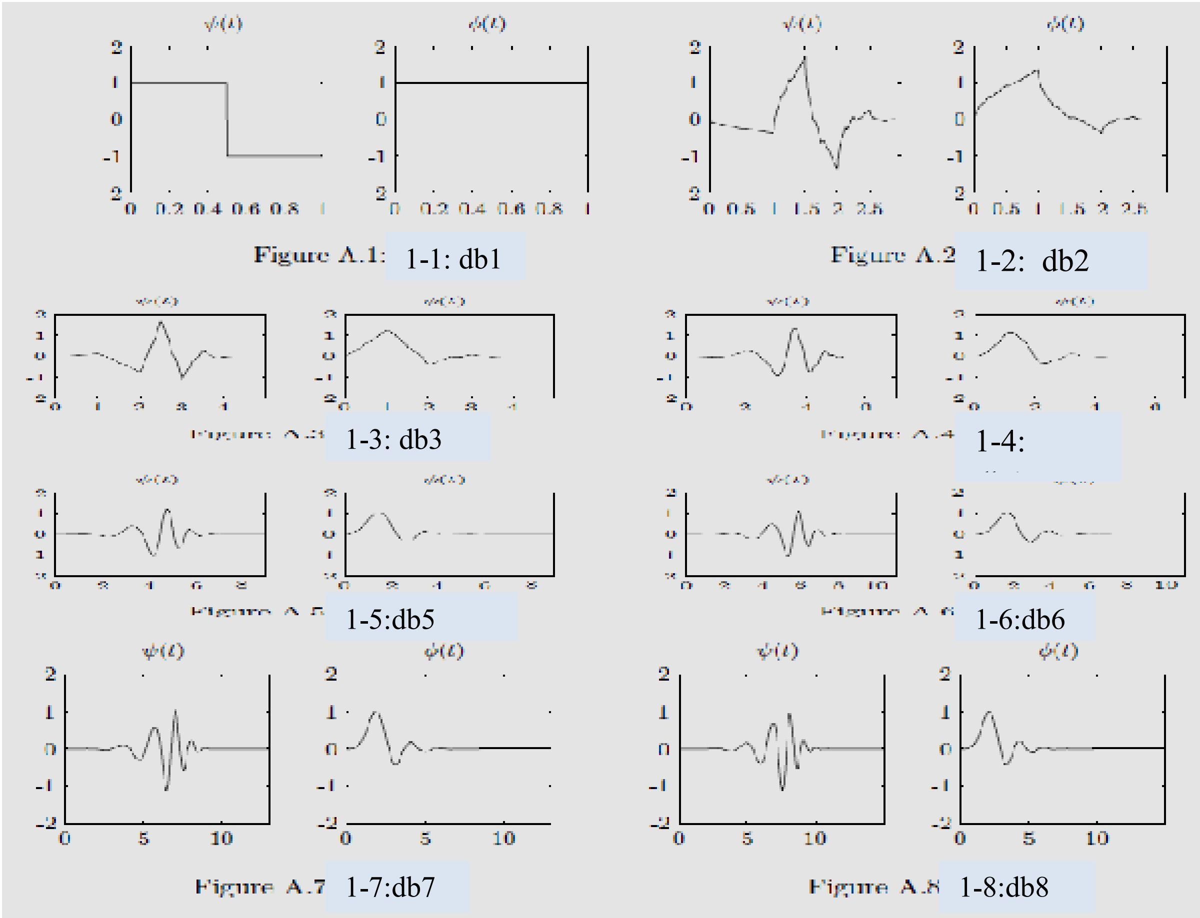


Figure 1. Daubechies Functions of orders 1 to 8.

Coiflets are also build by I. Daubechies on the request of R. Coifman. Coifman wavelets are orthogonal compactly supported wavelets with the highest number of vanishing moments for both the wavelet and scaling function for a given support with. The Coiflet wavelets are more symmetric and have more vanishing moments than the Daubechies wavelets [11]. The Coiflet wavelet and scaling functions with orders 1 up to 5 are shown in respectively Fig. 2-1 up to Fig. 2-5.

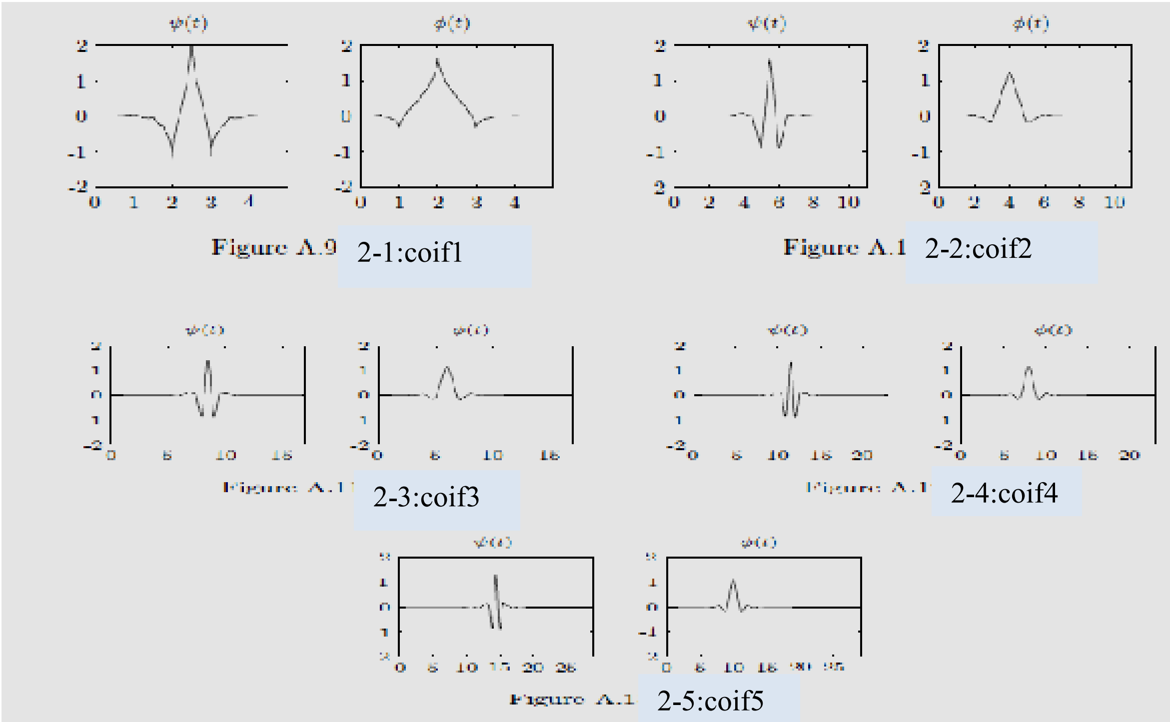


Figure 2. Coiflet wavelet functions of order 1 to 5

Symlets are also orthogonal and compactly supported wavelets, which are proposed by I. Daubechies as modifications to the db family. Symlets are near symmetric and have the least asymmetry. The associated scaling filters are near linear-phase filters. The properties of symlets are nearly the same as those of the db wavelets. The symlet wavelet and scaling functions for orders 2 up to 8 are shown in respectively Fig. 3-1 up to Fig. 3-8.

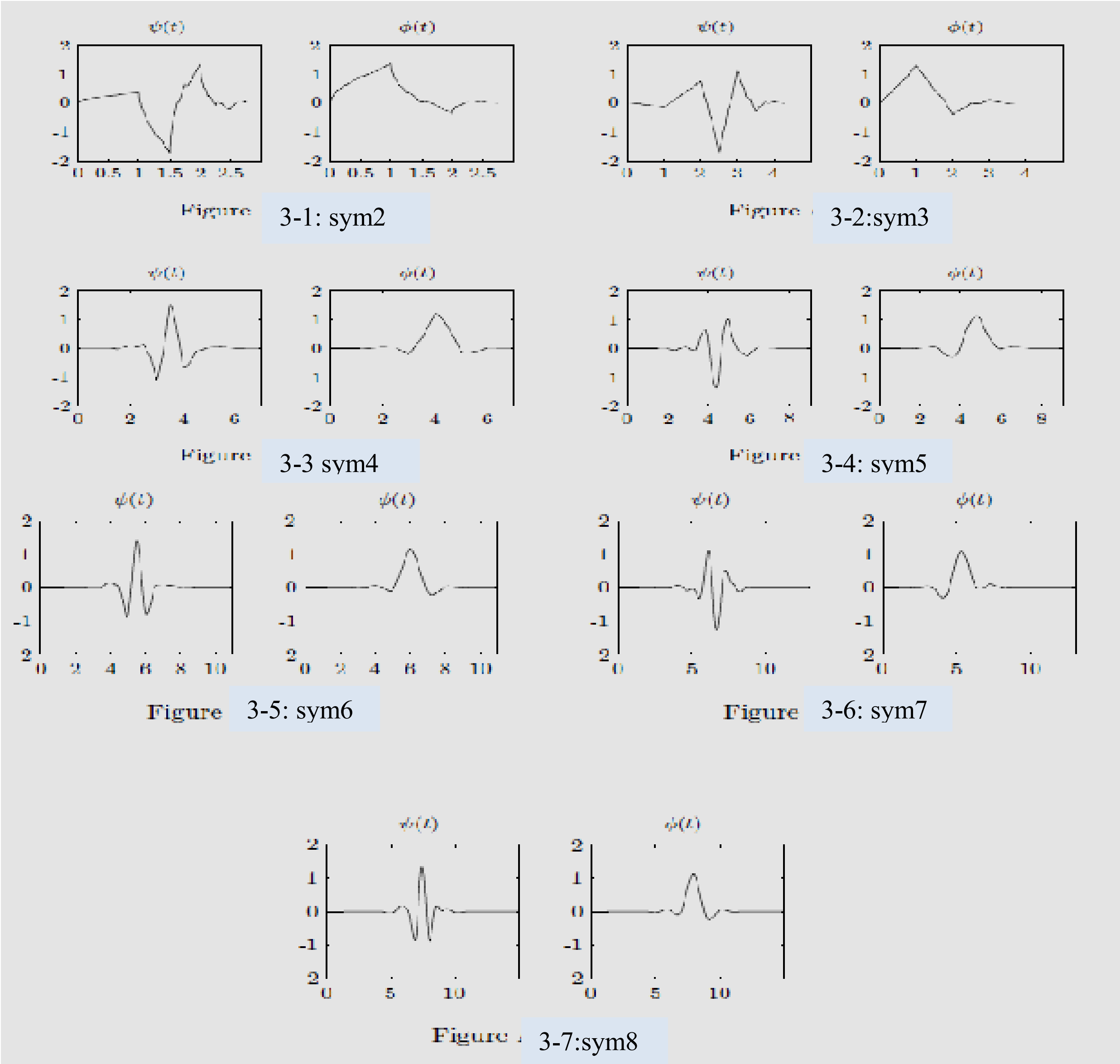


Figure 3-1: sym2 Figure 3-2:sym3

Figure 3-3 sym4 Figure 3-4: sym5

Figure 3-5: sym6 Figure 3-6: sym7

Figure 3-7:sym8

Figure 3. Symlet functions of order 1 to 8

3- The Denoising Algorithm

NighShrink algorithm states that, for each noisy wavelet coefficient w_{ij} to be shrinked, it incorporates a square neighbouring window B_{ij} centered at it. The neighbouring window size can be represented as LxL, where L is a positive odd number. Suppose $S_{ij}^2 = \sum_{k,l=B_{ij}} w_{ij}^2$, the NeighShrink shrinkage formula can be represented as

$$\hat{\theta} = w_{ij}\beta_{ij} \quad (2)$$

where $\hat{\theta}_{ij}$ is the estimator of the unknown noiseless coefficient.

$$\beta_{ij} = (1 - \lambda^2 / S_{ij}^2)_+ \quad (3)$$

and λ is the universal threshold. Different wavelet coefficient subbands are shrinked independently, but the threshold λ and neighbouring window size L keep unchanged in all subbands. The shortcoming of this method is that using the same universal threshold λ and neighbouring window size L in all subbands is suboptimal. In [8] NeighShrink method is improved by determining an optimal threshold and neighbouring window size for every wavelet subband using the Stein's unbiased risk estimate (SURE). For ease of notation, N_s noisy wavelet coefficients are arranged from subband s, $w_s = \{w_{ij}: \epsilon\, indices\ corresponding\ to\ subbands\}$, into the 1-D vector $w_s = \{w_n: n = 1, \ldots, N_s\}$.

Similarly, the N_s unknown noiseless coefficients are combined $\{\theta_{ij}: ij\epsilon\, indices\ corresponding\ to\ subbands\}$ from subband s into the corresponding 1-D vector $\theta_s = \{\theta_n: n = 1, \ldots, N_s\}$. Stein shows that, for almost any fixed estimator $\hat{\theta}$ based on the data w_s, the expected loss(i.e. risk) $E\{\|\hat{\theta}_s - \theta_s\|_2^2\}$ can be estimated unbiasedly. Usually, the noise standard deviation σ is set at 1, and we have that $E\left\{\|\hat{\theta}_s - \theta_s\|_2^2\right\} = N_s + E\{\|g(w_s)\|_2^2 + 2\nabla \cdot g(w_s)\}$ (4)

where $g(w_s) = \{g_n\}_{n=1}^{N_s} = \hat{\theta}_s - w_s, \nabla \cdot g \equiv \sum_n \frac{\partial g_n}{\partial w_n}$.

According to eq. (1), we have for the *n*th wavelet coefficient w_n:

$$g_n(w_n) = \widehat{\theta_n} - w_n = \begin{cases} -\frac{\lambda^2}{S_n^2} w_n & (\lambda < S_n \\ -w_n & otherwise \end{cases} \tag{5}$$

with

$$\frac{\partial g_n}{\partial w_n} = \begin{cases} -\lambda^2 \frac{S_n^2 - 2w_n^2}{S_n^4} & (\lambda < S_n \\ -1 & otherwise \end{cases} \tag{6}$$

$$\|g_n(w_n)\|_2^2 = \begin{cases} \frac{\lambda^4}{s_n^4} & (\lambda < s_n \\ w_n^2 & otherwise \end{cases} \tag{7}$$

The quantity

$$SURE(w_s, \lambda, L) = N_s + \sum_n \|g_n(w_n)\|_2^2 + 2\sum_n \frac{\partial g_n}{w_n} \tag{8}$$

is an unbiased estimate of the risk on subband s where L is the neighborhood window size (L is an odd number and greater than 1, for example, 3,5,7,9,etc.):

$E\left\{\|\hat{\theta}_s - \theta_s\|_2^2\right\} = E\{SURE(w_s, \lambda, L)\}$. The threshold $\lambda^s$and neighbouring window size L^son subband s which minimize SURE $(w_s, \lambda, L), accordingly,$

$$(\lambda^s, L^s) = argminSURE(w_s, \lambda, L) \tag{9}$$

Where $\lambda^s$and $L^s$are derived assuming the noise level$\sigma = 1$, for data with nonunit variance, the coefficients are standardized by an appropriate estimator $\hat{\sigma}$ before calculating the λ^s and L^s with Eq. (9). A good estimator for σ is the median of absolute deviation (MAD) using the highest level wavelet coefficients [2].

$$\hat{\sigma} = \frac{median(|w_s|)}{0.6745} (w_s \in subband\ HH)$$

(10)

The denoising algorithm with different wavelet families can be summarized as follows:-

1- A four-level 2-D discrete wavelet transform is performed to a noisy grey scale image contaminated with white Gaussian noise.

2- For each subband (HL_i, LH_i, HH_i), where $i = 1,2,3,4$. Determine the optimal threshold and neighbouring window size for the subbands using eq. (9).

3- Threshold the noisy subbands using NeighShrink rule.

4- Apply the inverse wavelet transform to get the denoised estimate image.

Fig. 4 shows the block diagram of the system.

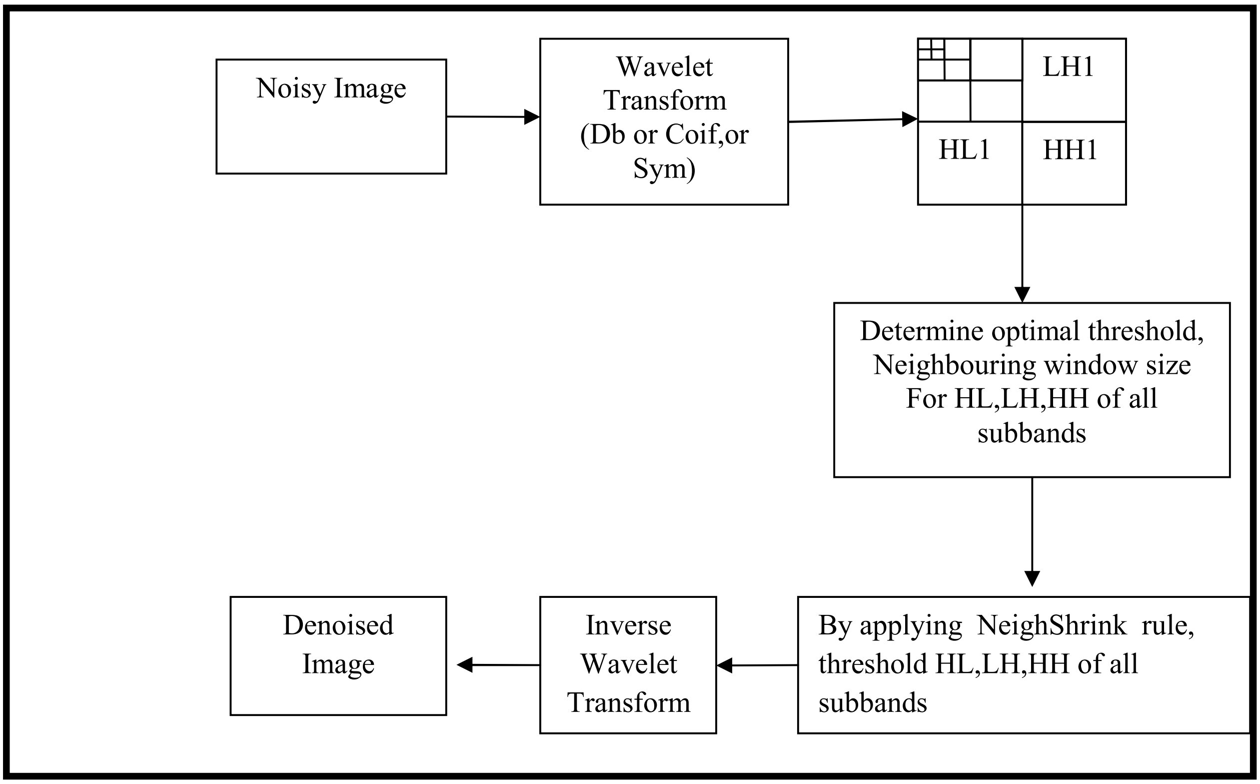


Figure 4. The block diagram of the denoising system

4- Experimental Results

The optimal threshold and neighbouring window method proposed by [8] is applied for Daubechies functions with order db1(Haar) and db4, Coiflet3 function, and the symlet8 function. The 512x512 grey scale standard test images, boat, Barbara, plane and mandrill were chosen as the experimental dataset Fig(5). They were contaminated with Gaussian random noise with standard deviations 10, 20, 30, 60, 80, and 100. In all wavelet subbands NeighShrink used the universal threshold $\lambda = \sigma\sqrt{2\log(512)} \approx 3.53\sigma$ and the default neighbouring window size 3x3 which is recommended by NeighShrink.

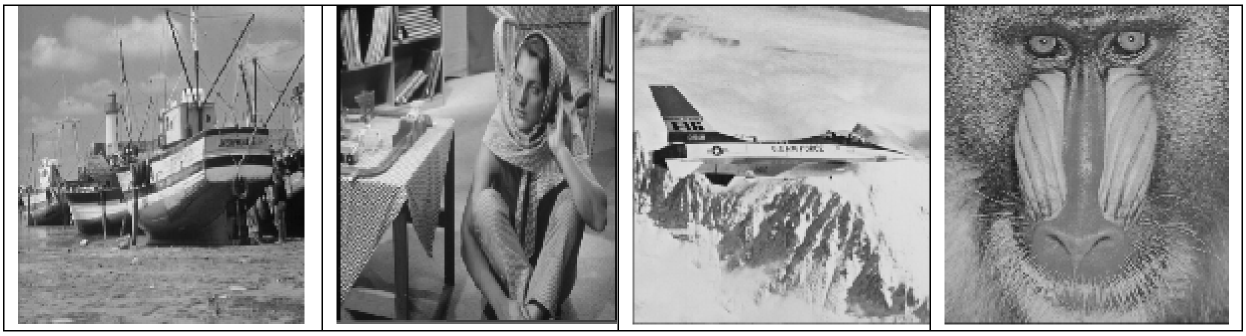


Figure 5. The original test images with 512x512 pixels

Fig. 6 shows the denoised images (boat, plane) for noise standard deviation 60,80 respectively, and the PSNR of applying wavelet functions (haar, coif3, sym8).

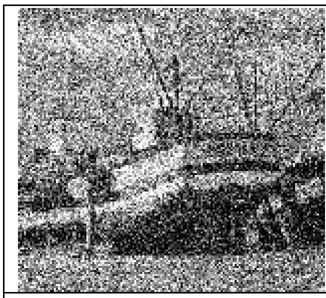
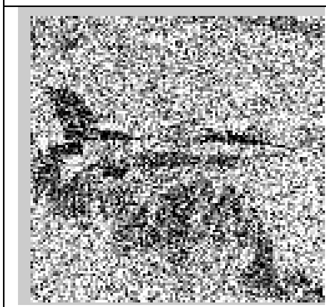
			
PSNR=12.56 dB	PSNR=24.09 dB	PSNR=25.3256 dB	PSNR=25.92 dB
			
PSNR=10.07dB	PSNR=21.235 dB	PSNR= 22.038 dB	PSNR=22.003 dB

Figure 6. Results of denoising process for haar, coiflet3, and sym8 wavelets

The experimental results are measured using peak signal-to-noise ratio (PSNR) in decibels (dB), which is defined as [9]:

$$PSNR = 10 * log_{10}\frac{255^2}{MSE} \ (dB) \tag{10}$$

where $MSE = \frac{1}{N^2}\sum_{i,j=1}^{N}(X(i,j) - \hat{X}(i,j))^2$, X is the original image, $\hat{X}$ is the estimate of X , and N^2 is the number of pixels. The denoised image is closer to the original one when PSNR is higher. Tables 1,2,3,4 shows the performance of the optimal threshold and neighbouring window method for Haar, Db4, Coiflet3, and Sym8 wavelet functions, with boat, Barbara, plane, mandrill test images.

Table 1. Denoising results for boat image

σ	Noisy boat image (PSNR)	Haar Wavelet PSNR	Db4 PSNR	Coiflet3 PSNR	Sym8 PSNR
10	28.13	32.79	34.00	34.16	33.99
20	22.11	28.96	30.26	30.41	30.24
30	18.58	26.97	28.23	28.37	28.23
60	12.56	24.09	25.19	25.33	25.19
80	10.06	23.06	24.07	24.20	24.03
100	8.13	22.31	23.19	23.35	23.15

Table 2. Denoising results for barbara image

σ	Noisy barbara image (PSNR)	Haar Wavelet PSNR	Db4 PSNR	Coiflet3 PSNR	Sym8 PSNR
10	28.13	31.58	32.73	32.95	33.02
20	22.11	27.46	28.79	29.03	29.09
30	18.58	25.38	26.70	26.95	27.00
60	12.56	22.47	23.70	23.87	23.87
80	10.06	21.52	22.61	22.76	22.75
100	8.13	20.87	21.85	21.95	21.93

Table 3. Denoising results for plane image

σ	Noisy plane image (PSNR)	Haar Wavelet PSNR	Db4 PSNR	Coiflet3 PSNR	Sym8 PSNR
10	28.13	33.34	34.31	34.41	34.42
20	22.11	29.34	30.63	30.74	30.76
30	18.58	27.16	28.64	28.70	28.67
60	12.56	23.85	25.43	25.40	25.37
80	10.06	22.80	24.11	24.14	24.13
100	8.13	22.03	23.19	23.22	23.17

Table 4. Denoising results for mandrill image

σ	Noisy plane image (PSNR)	Haar Wavelet PSNR	Db4 PSNR	Coiflet3 PSNR	Sym8 PSNR
10	28.13	29.89	30.24	30.30	30.95
20	22.11	25.61	26.11	26.20	26.19
30	18.58	23.52	24.08	24.12	24.16
60	12.56	20.76	21.35	21.38	21.39
80	10.06	19.96	20.48	20.52	20.51
100	8.13	19.45	19.94	20.00	19.94

From the results we can notice that the Haar wavelet function based denoising algorithm has the lowest values of PSNR while the Coiflet3 wavelet function has the highest values of PSNR and it's slightly better than Db4 and Sym8 wavelet functions. Also from the experimental results Haar wavelet function has a blocking effect on the denoising image.

6- Conclusion

In this paper wavelet transform families were applied to image denoising algorithm based on choosing an optimal threshold and neighbouring window. The wavelet families used were Haar, db4, coiflet3, and sym8 functions. The effect of using these functions were studied, the experimental results show that the response of the system was poorer with Haar, while it's approximately the same for other wavelet families.

References:

[1] Grace Chang,S., Bin Yu and Vattereli,M.," Adaptive Wavelet Thresholding for Image Denoising and Compression", *IEEE Trans. Image Processing*, vol. 9, pp. 1532-1546, Sept. 2000.

[2] Donoho, D.L., Johnstone, I.M.," Ideal spatial adaptation via wavelet shrinkage.Biometrika" 81, 425–455., 1994.

[3] Donoho, D.L., Johnstone, I.M.," Adapting to unknown smoothness via wavelet shrinkage". J. Amer. Statist. Assoc. 90, 1200–1224, 1995.

[4] Cai, T.T., Silverman, B.W., "Incorporating information on neighbouring coefficients into wavelet estimation". Sankhya, Ser. B 63, 127–148,2001.

[5] Chen, G.Y., Bui, T.D., Krzyzak, A., " Image denoising with neighbour dependency and customized wavelet and threshold". Pattern Recognition 38, 115–124 ,2005.

[6] Sendur, L., Selesnick, I.W.," Bivariate shrinkage with local variance estimation". IEEE Signal Process. Lett. 9, 438–441,2002.

[7] Luisier, F., Blu, T., Unser, M.," A new SURE approach to image denoising: interscale orthonormal wavelet thresholding". IEEE Trans. Image Process. 16, 593– 606,2007.

[8] Dengwen,Z., Wengang,C. " Image denoising with an optimal threshold and neighbouring window", Pattern Recognition Letters 29 (2008) 1694–1697.

[9] Sachin D Ruikar, Dharmpal D Doye, "Wavelet Based Image Denoising Technique", (IJACSA) International Journal of Advanced Computer Science and Applications,Vol. 2, No.3, March 2011.

[10] Craigmile, P.F. , Percival, D.B.," Wavelet-based trend detection and estimation. Technical report", University of Washington, Seattle, December ,2000.