2005 IEEE International
Symposium on Signal Processing
and Information Technology

An Efficient Parameter Selection Criterion for Image Denoising

Hamed Pirsiavash[1], Shohreh Kasaei[2], Farrokh Marvasti[1]
h_pirsiavash@ee.sharif.edu, skasaei@sharif.edu, marvasti@sharif.edu

[1]Multimedia Research Lab., Sharif University of Technology, Tehran, Iran
[2]Department of Computer Engineering, Sharif University of Technology, Tehran, Iran

Abstract - *The performance of most image denoising systems depends on some parameters which should be set carefully based on noise distribution and its variance. As in some applications noise characteristics are unknown, in this research, a criterion which its minimization leads to the best parameter set up is introduced. The proposed criterion is evaluated for the wavelet shrinkage image denoising algorithm using the cross validation procedure. The criterion is tested for some different values of thresholds, and the output leading to the minimum criterion value is selected as the final denoised output. The resulting outputs of our method and the previous threshold selection scheme for the wavelet shrinkage, i.e. the median absolute difference (MAD), are compared. The objective and subjective test results show the improved efficiency of the proposed denoising algorithm.*

Keywords – *image denoising, wavelet shrinkage, noise estimation, parameter selection.*

1. INTRODUCTION

Nowadays, by improving image acquisition systems, many types of cameras are available. Some of these cameras use very simple hardware in order to have low cost and to be embedded in other devices like mobile phones. Hence, the output images of these devices are noisy and poor. In addition, in most image processing systems, the taken image should be fed to some processing stages like compression and recognition. The parasitic noise in the input image could suffer the other processes and make them inefficient.

To overcome these shortcomings, many image denoising algorithms have been developed during recent years. For instance, Gaussian smoothing, neighborhood filtering, and wavelet shrinkage can be mentioned [1].

In general, all denoising methods have some parameters and thresholds which should be adjusted to gain the best performance. Generally, these parameters depend on the noise distribution and its variance. Most algorithms suppose the noise to have a white Gaussian distribution with a known variance. However, in practical situations, we have no information about the noise variance. Hence, another problem rises which is the parameter and threshold selection algorithm. During recent years, some researchers considered this problem and made some solutions [2, 3, and 4]. The *generalized cross validation* method is proposed by *Jansen et al*. for multiple wavelet threshold selection [2, 3]. They defined a criterion which its minimum roughly minimizes the *mean square error* (MSE), but their method works in some special conditions and as proved in [2], it works only for wavelet shrinkage with orthogonal transforms. In addition, as they mentioned in their paper, its output has low MSE, but it is not guaranteed to yield a good visual quality.

In this paper, assuming the additive noise to have an arbitrary distribution, a novel criterion for image denoising is introduced. The minimization of this criterion leads to near optimum parameter set for denoising purposes. In order to evaluate the performance of this criterion, it is applied for optimum parameter selection in a popular image denoising algorithm, the wavelet thresholding.

The layout of this paper is as follows: Section 2 introduces the proposed criterion and its efficiency in parameter selection. In section 3, wavelet shrinkage algorithm is described briefly. The experimental results and the performance comparison are presented in Section 4. Finally, Section 5 concludes the paper.

2. PROPOSED CRITERION FOR PARAMETER SELECTION

In image denoising algorithms with additive noise, the input image is assumed to be the summation of original image and an additive random noise. An important knowledge which is used in the proposed criterion is the independency of these two signals (the original image and the additive noise). Here, the aim of denoising algorithms is to remove the parasitic noise. In fact, the difference between the input and the output of the denoising stage is the estimated noise which has been removed (see Figure 1). Therefore, the distribution of the estimated noise should approach that of the additive noise.

The estimated noise for image denoising with two distinct parameter sets is shown in Figure 2. In this figure, the estimated noise is exaggerated to be shown clearly. It could be seen that there is a large similarity between the estimated noise and the original image, but this similarity in Figure 2(f) is less than that in Figure 2(e). It means that for an optimum image denoising algorithm, the correlation between the estimated noise and the output image which is an expectation of the original image should be minimized. This result is in agreement with the assumption of independency between the original image and the additive noise. Now, the correlation for each parameter set can be computed. Consequently, by minimizing that, the best parameter set for the denoising algorithm can be found.

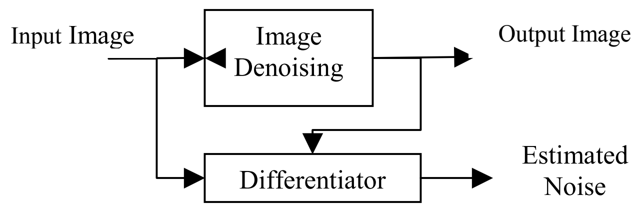


Fig. 1. Noise estimation

Fig. 2. Form left to right and top to bottom: (a) Original image(Lena), (b) noisy image (AWGN, sigma=15), (c) blurred denoised image, (d) denoised image using proper parameters, (e) estimated noise of (c), (f) estimated noise of (d). (The estimated noise is exaggerated to be seen clearly).

According to our knowledge, most denoising algorithms assume the original image to have more energy in lower frequency components compared to the noise. In addition, the additive noise usually has near flat spectrum (*i.e.* white noise). Hence, a huge energy of noise can be removed by removing the higher frequency components. But, we know that using these methods, the edge points which have higher frequency components will be blurred. As we know, because the human visualize system is more sensitive to the edges, the blurring effect will be perceived obviously. In this research, in order to adapt to the human visualize system, the criterion is altered. As can be seen in Figure 2, the edge points can be found in the estimated noise. Therefore, the correlation between the estimated noise and the edge map of the output image is used. Because the output image for some parameter sets has high amount of noise, the edge map should be extracted using a robust edge detection method; thus, here, Canny edge detector is used [5]. Then, the following value should be minimized.

$$C = Correlation(Estimated\ noise\ , output\ image\ edge) \quad (1)$$

With weak denoising parameters, the estimated noise approaches to a zero field and makes the correlation to have a low value. Consequently, the minimum correlation will be found for the weak denoising parameters. In order to suppress this defect, the estimated noise energy is used in the proposed criterion. The final criterion can be written as follows:

$$C_I = \frac{Correlation(Estimated\ noise\ ,\ output\ image\ edge)}{\log(Input\ image\ energy + estimated\ noise\ energy)} \quad (2)$$

The minimum of this criterion is found for the optimum parameters set. Some optimization methods like the genetic algorithms can be used to achieve the minimum, yet in this paper, showing the performance of the proposed criterion is the final goal; consequently, finding the minimum is performed using a simple method, i.e. cross validation. This criterion is computed for some parameter sets which are predefined multiplications of the parameter set of the MAD method, described in the next section, and the minimum value is chosen.

3. WAVELET SHRINKAGE

As the defined criterion should be evaluated and compared with the other parameter selection methods, in this research, a usual image denoising algorithm, i.e. the wavelet shrinkage is implemented. In this section, a brief description of this method is presented. The implementation results and details are discussed in Section 4.

Wavelet shrinkage is an efficient signal denoising algorithm introduced by *Donoho et al.* in [1, 6, and 7]. That method is based on the idea that the original image has large wavelet coefficients and the noise is distributed over all coefficients. Thus, by thresholding the small coefficients, the image will not be damaged although a large amount of noise energy will be removed. The hard thresholding is applied using:

$$\mathrm{HWT(x)} = \begin{cases} 0 & |x| \le T \\ x & |x| > T \end{cases} \quad (3)$$

where T is a predetermined threshold value. This basic idea causes some oscillations near the edges. As a result, they proposed soft thresholding method in which small wavelet coefficients are cancelled and the others are changed in order not to destroy the continuity in wavelet coefficients.

$$\mathrm{SWT(x)} = \begin{cases} 0 & |x| \le T \\ Sign(x) \cdot (|x| - T) & |x| > T \end{cases} \quad (4)$$

where *Sign(x)* denotes the *signum* function. Using this method, oscillations are suppressed [7].

In wavelet thresholding methods, the selection of thresholds for each resolution level is very important because according to the other denoising algorithms, wrong selection can make the output image blurred or noisy. Some threshold selection methods are introduced for Gaussian noise distributions with known variances. Three commonly used methods are the *universal, SURE,* and *MiniMax*. The mathematical details can be found in [6, 7]. For instance, universal method is as follows:

$$T = \hat{\sigma}_n \cdot \sqrt{\frac{2\log N}{N}} \quad (5)$$

where N is the number of data points and $\hat{\sigma}_n$ is the noise variance defined below. In most denoising algorithms, the *median of absolute difference* (MAD) is used for noise variance estimation [8].

$$\hat{\sigma}_n = \frac{MAD}{0.6745} = \frac{Median(|x - Median(x)|)}{0.6745} \quad (6)$$

This estimation yields to good results for Gaussian distributed noise. A typical wavelet shrinkage algorithm is shown in Figure 3.

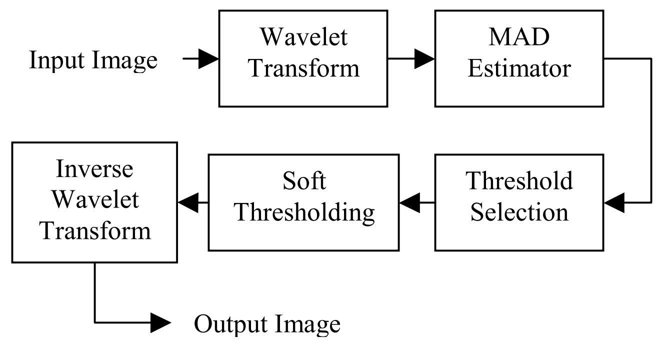


Fig. 3. Block diagram of a simple wavelet shrinkage denoising system.

4. EXPERIMENTAL RESULTS

In this section, the implementation results and the performance of the proposed algorithm when compared to other available approaches are presented. An efficient criterion is computed for several parameter sets in a denoising algorithm and the parameters leading to the minimum of the criterion are chosen as the best parameters for the input image and the related noise statistics. The proposed algorithm is implemented using Matlab package for wavelet shrinkage image denoising process.

As briefly discussed in Section 3, wavelet shrinkage is a powerful image denoising algorithm, and thus many researchers have proposed different modified versions of that algorithm. In this research, wavelet shrinkage is implemented in two resolution levels. Here, Daubechies wavelet with 6 tabs is used. The initial threshold for each subspace is chosen independently based on the MAD variance estimation and MiniMax threshold selection methods. Next, the criterion is computed for 11 different multiplications of this initial threshold set. The multiplications are chosen uniformly in the logarithmic scale in the range of $10^{-0.8}$ to $10^{1.2}$. Finally, the minimum is taken as the best solution. The MSE of this method and the MAD method for some standard images are listed in Table 1. The last column contains the best MSE among 11 tested threshold sets. As can be seen, the obtained denoising system (with selected variance) approaches the minimum MSE. The resulting PSNR and the calculated criterion for a sample image are plotted in Figure 4. As seen in this figure, choosing the minimum value of the proposed criterion matches the maximum PSNR that leads to the best parameter selection for denoising purposes. A sample output result is shown in Figure 5.

As seen in Table 1, for most tested images the MSE obtained from the proposed algorithm is less than that of the MAD method. Moreover, for some cases the obtained MSE by our algorithm is close to the minimum available MSE.

As the MSE is not the best measurement for performance analysis in image processing systems, the outputs should be examined in a subjective test as well. The results of the subjective test among 20 boys and girls are presented in Table 2. Some particular cases in which the MSEs of our method are high are examined in this test. For instance, in the 21st row of Table 1, the resulting MSE is higher than that of the MAD method. The outputs of the 21st case are shown in Figure 6. It is obvious that the output of our method is subjectively better than the MAD output, which proves that both objective and subjective tests should be run. In fact, because the criterion uses the edge map, our algorithm leads to less defects in the edge areas and thus results in higher subjective performance; although it may have a higher MSE. Another result obtained from these implementations is that for lower input noise variances, our method performs much better than the MAD method. Because for images with a low level of noise, after denoising the edge map can be extracted more efficiently, and thus the proposed algorithm can better calculate the minimum that matches the maximum of the PSNR. As another result, for images with small size (about 256x256), our method performs better than the MAD and MiniMax methods, because our method is less directly dependent to the statistics of the images. This fact motivated us to examine this method in spatially adaptive wavelet shrinkage algorithms introduced in [9].

5. CONCLUSION

In this paper, an efficient criterion for performance analysis of denoising systems is introduced. It is shown that using a cross validation procedure, we can adjust the system parameters to achieve a better performance. This criterion is examined for wavelet shrinkage as a common denoising algorithm. According to the results, the obtained subjective tests show the superiority of the proposed algorithm when compared to the MAD approach. Another important advantage of this method is its independency on the noise distribution and its variance. As mentioned above, the algorithm performs even better for images with lower noise variances.

Table 1. MSE of our method in comparison with MAD and minimum available MSE.

	Image name	Noise standard deviation	MSE of proposed method	MSE of MAD method	Min available MSE
1	Barbara512	5	18.79	112.18	18.79
2	Barbara512	10	56.22	173.19	56.22
3	Barbara512	15	100.89	222.49	100.89
4	Barbara512	20	201.52	260.27	152.77
5	Barbara512	25	228.89	292.18	197.45
6	Barbara256	5	22.66	162.87	22.66
7	Barbara256	10	70.74	237.11	65.52
8	Barbara256	15	131.45	294.70	121.27
9	Barbara256	20	190.61	345.20	190.61
10	Lena512	5	12.84	35.57	12.84
11	Lena512	10	32.87	55.13	32.87
12	Lena512	15	58.44	73.91	55.79
13	Lena512	20	84.00	92.65	78.10
14	Lena512	25	98.74	110.67	98.74
15	Lena256	5	17.35	79.24	17.35
16	Lena256	15	95.65	153.05	88.24
17	Boat	5	18.45	80.26	18.45
18	Boat	10	54.47	107.66	48.90
19	Boat	15	113.38	133.26	82.43
20	Boat	15	81.59	133.68	81.59
21	Boat	20	192.18	157.22	114.07
22	Peppers	5	15.18	38.55	15.18
23	Peppers	10	49.97	55.90	35.56
	Average	**13.04**	**84.65**	**148.13**	**74.19**

Table 2. Subjective test results among 20 boys and girls (score 5 is assigned to the original image).

No.	Image name	Noise standard deviation	Proposed method	MAD method
1	Barbara512	25	3.30	2.62
2	Barbara256	20	3.47	2.44
3	Lena256	15	3.49	2.99
4	Boat	15	4.12	3.19
	Average	**18.75**	**3.59**	**2.81**

ACKNOWLEDGEMENT

The authors would like to acknowledge the support of Iran Telecommunication Research Center (ITRC) for this project.

REFERENCES

[1] A. Buades, B. Coll, and J. M. Morel, “On image denoising methods,” CMLA Preprint (http://www.cmla.ens-cachan.fr/Cmla/), to appear in *SIAM Multiscale Modeling and Simulation*, 2004.

[2] M. Jansen, M. Malfait and A. bultheel, "Generalized cross validation for wavelet thresholding," *Signal Processing*, Vol. 56, pp. 33-34, Jan 1997.

[3] M. Jansen and A. Bultheel, "Multiple wavelet threshold estimation by generalized cross validation for images with correlated noise," *IEEE Tarns. on Image Processing*, Vol. 8, No. 7, pp. 947-953, Jul 1999.

[4] Tai-Chiu Hsung and D. Pak-Kong Lun, "Generalized cross validation for multiwavelet shrinkage," *IEEE Signal Processing Letters*, Vol.11, No. 6, pp. 549-552, Jun 2004.

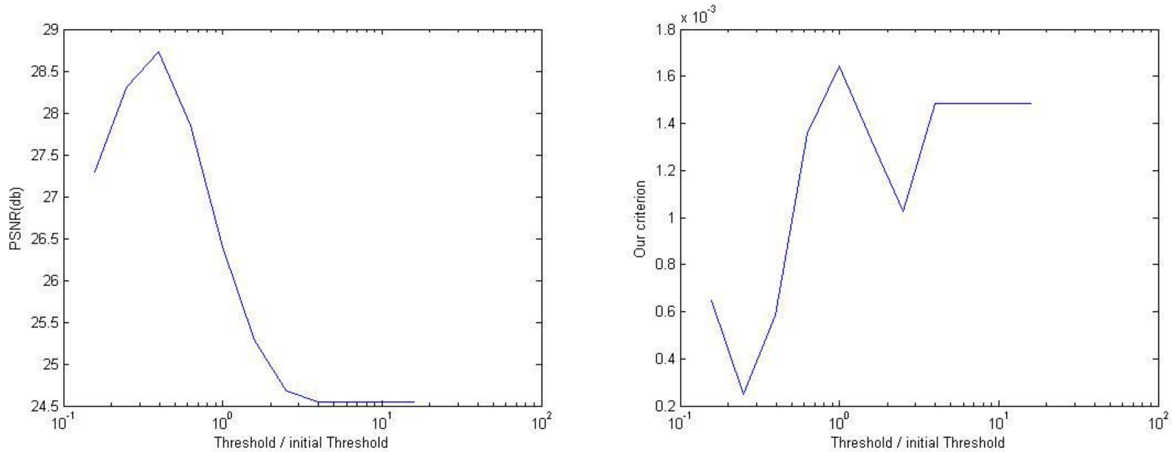
[5] Canny and John, "A Computational Approach to Edge Detection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. PAMI-8, No. 6, pp. 679-698, 1986.

[6] D. Donoho and I. Johnstone, "Ideal spatial adaptation via wavelet shrinkage," *Biomedika*, vol. 81, pp. 425–455, 1994.

[7] D. Donoho, "Denoising by soft-thresholding," *IEEE Trans. on Information Theory*, Vol. 41, pp.613-627, 1995.

[8] P. J. Rousseeuw and A. M. Leroy, *Robust Regression and Outlier Detection*, New York: Wiley, 1987.

[9] S. Grace Chang, Bin yu and Martin Vetterli, "Spatially adaptive wavelet thresholding with context modeling for image denoising," *IEEE Trans. on Image Processing*, vol. 9, No. 9, pp. 1522-1531, Sept 2000.



(a) (b)

Fig. 4. Results for Lena256 (with Gaussian noise, standard deviation=15) (a) PSNR vs. threshold ratio, (b) proposed criterion vs. threshold ratio. (Threshold is the multiplication of the threshold ratio and the MAD threshold, *i.e.* setting threshold ratio equal to one leads to the MAD method).

Fig. 5. From left to right and top to bottom: (a) Original image (Barbara256), (b) noisy image (Gaussian, standard deviation=15), (c) MAD output (MSE=295), (d) output of our method (MSE=131).

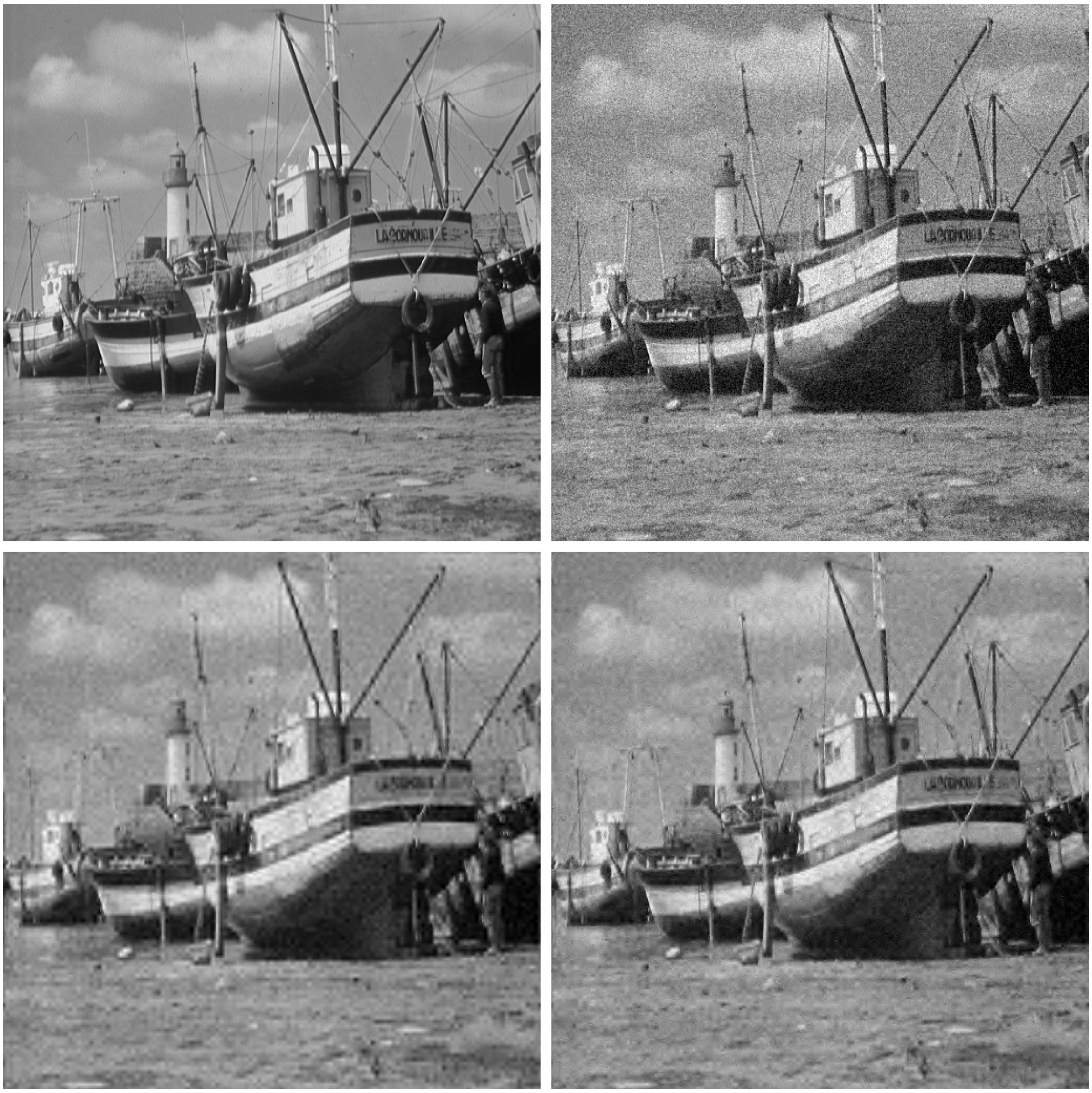


Fig. 6. From left to right and top to bottom: (a) Original image (fishing boat), (b) noisy image (AWGN, standard deviation=20), (c) MAD output (MSE=157), (d) output of our method (MSE=192).