AN ADAPTIVE THRESHOLDING APPROACH FOR IMAGE DENOISING USING REDUNDANT REPRESENTATIONS

*Zahra Sadeghipour, Massoud Babaie-Zadeh**

Electrical Engineering Department
Sharif University of Technology
Tehran, Iran

Christian Jutten

GIPSA-lab, Grenoble, France
and Institut Universitaire de France

ABSTRACT

A frequently used approach for denoising is the shrinkage of coefficients of the noisy signal representation in a transform domain. Although the use of shrinkage is optimal for Gaussian white noise with complete and unitary transforms, it has already been shown that shrinkage has promising results even with redundant transforms. In this paper, we propose using adaptive thresholding of redundant representations of the noisy image for image denoising. In the proposed thresholding scheme, a different threshold is used for each representation coefficient of the noisy image in an overcomplete transform. In this method, each threshold is automatically set based on statistical properties of the noise in the redundant transform domain. In our algorithm, adaptive thresholding is applied to redundant representations of noisy image blocks. Simulation results show that our method achieves the state-of-the-art denoising performance.

Index Terms- Image denoising, redundant representation, adaptive thresholding.

1. INTRODUCTION

In this paper, we address the image denoising problem, where an ideal image is measured or stored in presence of the noise. Consider the following model:

$$y = x + \nu \qquad (1)$$

where the noise free image is vectorized to form the vector x, ν denotes the noise assumed to be zero mean white Gaussian, and y is the measured image. The goal of image denoising is to remove noise from y to recover the original image, x.

Many different approaches such as spatial adaptive filters, morphological analysis, transform domain methods and more have been already studied for image denoising [1]. Here we focus on transform domain methods. These methods consist of three main steps: first applying the forward transform on the noisy image to represent the image in a transform domain, then applying the shrinkage operator on representation coefficients to attenuate the noise, finally applying the inverse transform on thresholded coefficients to obtain the denoised image. In 1995, Donoho et al. [2] proposed applying the shrinkage to noisy wavelet coefficients for signal denoising. The wavelet transform which has been used in [2] is complete and unitary, however it has been first experimentally [3], [4], [5] and then theoretically [6] shown that coefficient shrinkage has promising results in signal denoising even with *redundant* transforms. In recent years there are many research interests towards overcomplete transforms for image processing tasks (see [7] and the references therein). The problem of representing a signal x over an overcomplete transform can be expressed as follows:

$$x^{n \times 1} = D^{n \times m} . s^{m \times 1} \qquad m > n \qquad (2)$$

where the matrix D is usually called *dictionary*, s is the vector of transform coefficients and the superscripts denote the dimension. As was said, Elad in [6] has theoretically shown why coefficient shrinkage works with redundant transforms, however in [6], he has not mentioned the problem of setting the threshold automatically. Similarly in some other papers who have used coefficient shrinkage in the redundant transform, the automatic selection of the threshold has not been addressed (see [6] and the references therein). In fact up to our best knowledge, among these papers this problem has been addressed only in [7]. Authors of [7] assumed that representation coefficients of the noisy image in the redundant transform domain are independent Laplacian variables, and set the threshold based on this assumption. As choosing a proper threshold is an important issue in the shrinkage, in this paper we consider the problem of setting a suitable threshold automatically where the transform is re-

dundant. *Indeed, we propose an algorithm which adaptively thresholds representation coefficients of the noisy image in a redundant transform domain.* In the proposed thresholding scheme different thresholds are used for each transform coefficient of the noisy image and these thresholds are calculated automatically. As was said, it has been also proposed in [7] to set thresholds automatically in the redundant transform domain, but we use a different approach to set thresholds and simulation results show that our algorithm outperforms the denoising algorithm which has been proposed in [7]. Moreover it should be noted that authors of [7] have proposed an iterative method for signal denoising by imposing the *sparsity* constraint on the signal representation. To find a sparse representation of a signal, one needs to minimize the ℓ^0 norm of a signal representation. The ℓ^0 norm of a vector is the number of its nonzero components. It has already been proved [8], that minimizing the ℓ^0 norm of a vector is equivalent to minimization of its ℓ^1 norm when a highly sparse representation exists, so a family of algorithms called Basis Pursuit (BP) [9] tries to find the sparse representation by minimizing the ℓ^1 norm. The algorithm which has been proposed in [7] solves the Basis Pursuit De-Noising (BPDN) problem [9]. Solving the BPDN problem, denoises a signal by finding a sparse representation of the noisy image over a redundant transform. However, contrary to [7], our algorithm is not based on sparse representation over a redundant transform. It is simply based on the minimum ℓ^2 norm representation of the noisy image block. Since calculation the minimum ℓ^2 norm representation can be done highly faster than finding the sparse representation, our algorithm is a faster denoising algorithm.

In this paper we set thresholds based on statistical properties of the noise in the redundant transfrom domain. Donoho et al. in [2] have also proposed to set thresholds for shrinkage based on statistical properties of the noise in the wavelet domain. The main difference between our work and the algorithm proposed in [2] is that they used complete transforms while, we are using overcomplete transforms. In effect, the authors of [2] have derived statistical properties of the noise in the complete and unitary wavelet domain and their results cannot be used when the used transform for denoising is overcomplete. In this paper, we derive mathematical expressions which relate statistical properties of the noise in the redundant transform domain to corresponding properties of the noise in the spatial domain. It should be emphasized that we assume zero mean white Gaussian noise with known variance is added to the image as assumed in [7], [10], [11].

To summarize, our algorithm like any other redundant transform domain method consists of three main steps: finding transform coefficients of the noisy image block in the redundant transform, thresholding transform coefficients and applying the inverse transform on thresholded coefficients to obtain the denoised image. The novelty of this paper is using adaptive thresholding which uses a different threshold for each representation coefficient of the noisy image. The used thresholds are automatically set based on statistical parameters of the noise in the redundant transform domain.

This paper is organized as follows. In Section 2, we state statistical properties of the noise in a redundant transform domain based on corresponding properties in the spatial domain. The proposed algorithm for image denoising is developed in Section 3. In Section 4, we show some simulation results to demonstrate the effectiveness of the proposed algorithm.

2. STATISTICAL PROPERTIES OF NOISE IN REDUNDANT TRANSFORM DOMAIN

In this section, we derive statistical parameters of the noise in the redundant transform domain based on corresponding properties of the noise in the spatial domain. To this aim, consider the following equation:

$$\boldsymbol{\nu}^{n \times 1} = \boldsymbol{D}^{n \times m}.\boldsymbol{\alpha}^{m \times 1} \qquad m > n \qquad (3)$$

Here $\boldsymbol{\nu}$ represents noise in the spatial domain. $\boldsymbol{\nu}$ is a vector of zero mean white Gaussian random variables with variance σ^2 and $\boldsymbol{\alpha}$ is its representation in the transform domain. Among the infinite number of solutions of (3), we consider the minimum ℓ^2 norm solution, so:

$$\begin{aligned} \tilde{\boldsymbol{\alpha}} &= \boldsymbol{D}^\dagger.\boldsymbol{\nu} \\ &= \boldsymbol{B}.\boldsymbol{\nu} \end{aligned} \qquad (4)$$

where $\boldsymbol{D}^\dagger$ is the Moore-Penrose pseudo inverse of $\boldsymbol{D}$, $\boldsymbol{B} \triangleq \boldsymbol{D}^\dagger$ and $\tilde{\boldsymbol{\alpha}}$ is the minimum ℓ^2 norm solution of (3). Let $\boldsymbol{b}_i^T$ be a column vector which denotes the ith row of $\boldsymbol{B}$, the ith entry of $\tilde{\boldsymbol{\alpha}}$ can be written as:

$$\tilde{\alpha}_i = \boldsymbol{b}_i^T.\boldsymbol{\nu} \qquad (5)$$

Now we can state the following expressions about the statistical properties of $\tilde{\boldsymbol{\alpha}}$.

1. Gaussianity: Clearly $\tilde{\alpha}_i$ is a linear combination of entries of $\boldsymbol{\nu}$, which are assumed to be Gaussian random variables, so $\tilde{\alpha}_i$ is a Gaussian random variable.

2. Expected value of $\tilde{\alpha}_i$: The elements of $\boldsymbol{\nu}$ had been assumed to be zero mean variables. So applying the mathematical expectation on both sides of (5) yields:

$$\begin{aligned} E\{\tilde{\alpha}_i\} &= E\{\boldsymbol{b}_i^T.\boldsymbol{\nu}\} \\ &= \boldsymbol{b}_i^T.E\{\boldsymbol{\nu}\} = 0 \end{aligned} \qquad (6)$$

3. Expected value of $\tilde{\alpha}_i^2$: The entries of $\boldsymbol{\nu}$ had been assumed to be independent random variables with variance σ^2, so:

$$E\{\boldsymbol{\nu}\boldsymbol{\nu}^T\} = \sigma^2\boldsymbol{I} \qquad (7)$$

Combining (5) and (7) yields:

$$E\{\tilde{\alpha}_i^2\} = E\{(\boldsymbol{b}_i^T.\boldsymbol{\nu})(\boldsymbol{\nu}^T.\boldsymbol{b}_i)\} = \boldsymbol{b}_i^T E\{\boldsymbol{\nu}\boldsymbol{\nu}^T\}\boldsymbol{b}_i$$
$$= \boldsymbol{b}_i^T \sigma^2 \boldsymbol{I} \boldsymbol{b}_i = \sigma^2 \|\boldsymbol{b}_i\|_2^2 \tag{8}$$

4. Covariance of the entries of $\tilde{\alpha}$: Based on (5) the covariance of the entries of $\tilde{\alpha}$ can be written as:

$$E\{\tilde{\alpha}_i \tilde{\alpha}_j\} = E\{(\boldsymbol{b}_i^T.\boldsymbol{\nu})(\boldsymbol{\nu}^T.\boldsymbol{b}_j)\}$$
$$= \boldsymbol{b}_i^T \sigma^2 \boldsymbol{I} \boldsymbol{b}_j = \sigma^2 \boldsymbol{b}_i^T.\boldsymbol{b}_j \tag{9}$$

So the covariance of the elements of $\tilde{\alpha}$ depends on correlation between the rows of matrix $\boldsymbol{B}$. Matrix $\boldsymbol{D}$ is overcomplete therefore it cannot be orthogonal, as a consequence rows of $\boldsymbol{B}$ are not orthogonal to each other. It means that the elements of $\tilde{\alpha}$ are dependent.

We can summarize all of the above results to say: elements of $\tilde{\alpha}$ are zero mean colored Gaussian variables. The variance of the ith element of $\tilde{\alpha}$ is $\sigma^2 \|\boldsymbol{b}_i\|_2^2$, where $\boldsymbol{b}_i^T$ is the ith row of the pseudo inverse matrix of $\boldsymbol{D}$. In the next section, we use these results to set thresholds for adaptive thresholding.

3. ADAPTIVE THRESHOLDING FOR IMAGE DENOISING

In this section, we first explain the whole denoising process for a noisy image block, then we describe the details of the proposed algorithm. Denoising each image block consists of three main steps: finding transform coefficients of the noisy image block, coefficient shrinkage to attenuate the noise and applying the inverse transform on thresholded coefficients to obtain the denoised image block. Let $\boldsymbol{y}$ be a block of size $\sqrt{n} \times \sqrt{n}$ in the noisy image and $\boldsymbol{x}$ be the corresponding block in the noise free image. If we represent $\boldsymbol{y}$ and $\boldsymbol{x}$ in vector form, we have:

$$\boldsymbol{y}^{n \times 1} = \boldsymbol{x}^{n \times 1} + \boldsymbol{\nu}^{n \times 1}, \tag{10}$$

where $\boldsymbol{\nu}$ is the zero mean white Gaussian noise with variance σ^2. Consider the following equation:

$$\boldsymbol{y}^{n \times 1} = \boldsymbol{D}^{n \times m}.\boldsymbol{s}_1^{m \times 1} \qquad m > n \tag{11}$$

where $\boldsymbol{D}$ is the dictionary of the redundant transform and $\boldsymbol{s}_1$ is the noisy image block representation in the transform domain. Among the infinite number of solutions of (12), we choose the minimum ℓ^2 norm solution, so:

$$\boldsymbol{s}_1 = \boldsymbol{D}^\dagger.\boldsymbol{y} \tag{12}$$

In the next step, to attenuate noise, shrinkage operator is applied to transform coefficients. We use hard thresholding for shrinkage because it faithfully preserves the fine
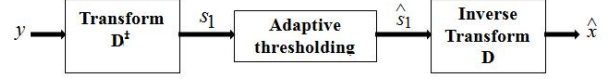
Fig. 1. A block diagram of the transform domain method used in this paper for image denoising

details of the image. We propose to use adaptive thresholding which uses different threshold for each representation coefficient of $\boldsymbol{y}$. Let s_{1_i} be the ith entry of $\boldsymbol{s}_1$ and t_i be its corresponding threshold, the thresholded coefficient $\widehat{s}_{1_i}$, is then expressed as follows:

$$\widehat{s}_{1_i} = \begin{cases} s_{1_i}, & \text{if} \quad |s_{1_i}| > t_i \\ 0, & \text{if} \quad |s_{1_i}| < t_i \end{cases} \tag{13}$$

We will explain later, how the thresholds for adaptive thresholding are set. Finally the corresponding block of the denoised image is obtained by applying the inverse transform on $\widehat{\boldsymbol{s}_1}$:

$$\widehat{\boldsymbol{x}} = \boldsymbol{D}.\widehat{\boldsymbol{s}_1} \tag{14}$$

Figure 1 illustrates the described denoising algorithm.

In the following we explain how adaptive thresholds for coefficient shrinkage are set to remove the noise effectively. To set the proper threshold for signal denoising, Donoho et al. in [2] used the following theorem.

Theorem 1. : If $\{z_i\}$ is a sequence of m zero mean white Gaussian random variables with variance σ^2 then:

$$Prob\{max|z_i| > \sigma\sqrt{2\log m}\} \to 0, \quad as \quad m \to \infty.$$

It means that with a high confidence all z_i's are lower than $\sigma\sqrt{2\log m}$. So if a signal of length m is contaminated by white Gaussian noise with variance σ^2, hard thresholding with threshold $\sigma\sqrt{2\log m}$ highly attenuates the noise.

Since in our algorithm hard thresholding is applied to representation coefficients, knowing statistical properties of the noise in the transform domain is necessary to choose a proper threshold using Theorem 1. Combining (4), (10) and (12) yields:

$$\boldsymbol{s}_1 = \boldsymbol{D}^\dagger.\boldsymbol{y} = \boldsymbol{D}^\dagger.(\boldsymbol{x} + \boldsymbol{\nu})$$
$$= \boldsymbol{D}^\dagger.\boldsymbol{x} + \tilde{\alpha} \tag{15}$$

So if we use the minimum ℓ^2 norm representation of the noisy image block, the noise representation in the transform domain is $\tilde{\alpha}$, which its statistical properties are stated in Section 2. As stated in Section 2, elements of the representation vector of zero mean white Gaussian noise in a redundant transform domain are zero mean colored Gaussian variables. It has been shown that variances of different components of the noise representation are different, so a fixed

threshold for all of representation coefficients of the noisy image cannot be used. Therefore, we propose using adaptive thresholding of (13) where t_i depends on the variance of the ith coefficient of the noise in the transform domain. Using Theorem 1, we propose to choose t_i as follows:

$$t_i = \sigma_i \sqrt{2 \log m} \tag{16}$$

where m is the length of representation vector of the noisy image block in a transform domain and σ_i is the standard deviation of ith representation coefficient of the noise. Combining (8) and (16) yields:

$$t_i = \|\boldsymbol{b}_i\|_2 \sigma \sqrt{2 \log m} \tag{17}$$

where $\boldsymbol{b}_i^T$ is the ith row of the pseudo inverse matrix of $\boldsymbol{D}$. It seems necessary to explain that, components of the noise in the redundant transform domain satisfy all of conditions of Theorem 1 except that they are dependent as stated in Section 2. However we experimentally find it useful to set thresholds as (17) even if noise coefficients in the transform domain are not independent.

It should be noted that, in this paper we do not mention the problem of choosing the optimal dictionary for image denoising, and simply use the redundant DCT dictionary which we find it experimentally useful for denoising natural images. The redundant DCT dictionary of size $n \times k$, (where $k > n$) is formed simply by n first rows of the complete DCT dictionary of size $k \times k$. The pseudo code of the proposed algorithm is shown in Fig. 2.

4. SIMULATION RESULTS

In this section we present some simulation results to demonstrate the performance of our algorithm. In all of the experiments zero mean white Gaussian noise is added to test images unless in experiments in which the effect of the noise type is investigated. Every result of our algorithm is an average over 20 experiments, having different realizations of the noise.

First, we compare the performance of our algorithm with the algorithm presented in [7]. We do not repeat the simulation for this algorithm and only mention the results which have been reported in [7]. As was said, authors of [7] have used shrinkage for image denoising. They have set a proper threshold for shrinkage by assuming a Laplacian model for representation coefficients of the noisy image. We use the overcomplete DCT matrix of size 64×256 as the dictionary of the redundant transform in our algorithm. This dictionary is used for denoising an image block of size 8×8 pixels. Table 1 reports denoising results of the algorithm proposed in [7] and our algorithm in terms of Peak Signal to Noise Ratio (PSNR) which is defined as follows:

$$\text{PSNR} = 10 \log \frac{255^2}{\frac{1}{N} \sum_{i=1}^{N} |x_i - \widehat{x}_i|^2}$$

- *Given parameters*:
  1. σ is the standard deviation of the noise in the spatial domain (assumed to be known).
  2. $\boldsymbol{D} \in \mathbb{R}^{n \times m}$ is an overcomplete dictionary (assumed to be given).

- *Initialization*:
  1. Compute the pseudo inverse of $\boldsymbol{D}$ and call it $\boldsymbol{B}$.
  2. Compute the adaptive threshold as the following: $t_i = \|\boldsymbol{b}_i\|_2 \sigma \sqrt{2 \log m}$, here $\boldsymbol{b}_i^T$ is the ith row of $\boldsymbol{B}$.

- *Repeat the following steps until all of noisy image blocks are denoised*:
  1. Consider a noisy image block of size $\sqrt{n} \times \sqrt{n}$ and vectorize it to form $\boldsymbol{y}$.
  2. Find the representation vector of $\boldsymbol{y}$, $\boldsymbol{s_1}$ in the transform domain as the following:
     $$\boldsymbol{s_1} = \boldsymbol{D}^\dagger . \boldsymbol{y}$$
  3. Threshold each element of $\boldsymbol{s_1}$ as follows:
     $$\widehat{s}_{1_i} = \begin{cases} s_{1_i} & |s_{1_i}| > t_i \\ 0 & |s_{1_i}| < t_i \end{cases}$$
  4. Obtain the denoised block as follows:
     $$\widehat{\boldsymbol{x}} = \boldsymbol{D} . \widehat{\boldsymbol{s_1}}$$
  5. Consider the next block and go to step 1.

Fig. 2. The adaptive thresholding algorithm for image denoising

---

where $\boldsymbol{x}$ is the ideal image of size $\sqrt{N} \times \sqrt{N}$ and $\widehat{\boldsymbol{x}}$ is the denoised image. As can be seen in Table 1 our algorithm outperforms the BPDN algorithm which is presented in [7] for all test images. Our algorithm outperforms about 1 dB in the cases of "Peppers" and "Lena" images and it outperforms about 0.8 dB in the case of "Barbara" image.

We also compare our algorithm with the OMP algorithm. Using OMP algorithm for image denoising is explained in [10]. Note that this algorithm does not explicitly use shrinkage for image denoising. To simulate the OMP algorithm we used MATLAB codes available at `http://www.cs.technion.ac.il/~elad/Various/KSVD_Matlab_ToolBox.zip`. In both of proposed and OMP algorithms we use the redundant DCT dictionary of size 64×256. The simulation results are reported in Table 2 in terms of PSNR. As can be seen in Table 2 the average output PSNR of two algorithms are close to each other. In the case of high noise power,

Table 1. Denoising PSNR results in decibels of the algorithm proposed in [7] (BPDN) and our algorithm (adaptive thresholding). Noise variance = 20

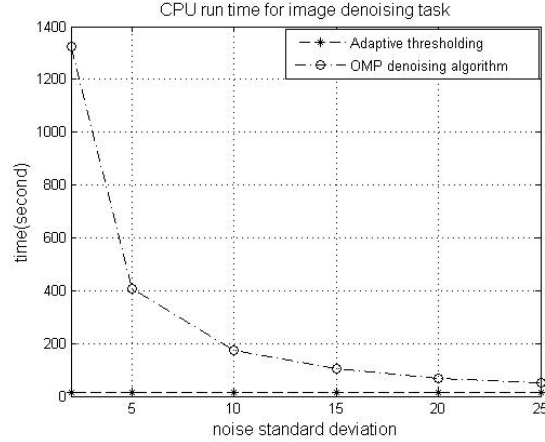| algorithm	Peppers	Lena	Barbara
BPDN	29.21	31.49	29.61
Adaptive thresholding	30.18	32.50	30.43

Table 2. Denoising PSNR results in decibels. For each noise level two results are presented. The top result is the output PSNR of the proposed algorithm and the bottom one is the result of the denoising algorithm proposed in [10].

σ/PSNR	Lena	Barb	House	Peppers	Average
2/42.11	44.25	43.67	44.44	43.36	43.93
	44.25	43.61	44.40	43.33	43.90
5/34.16	38.84	37.15	39.09	37.66	38.18
	38.86	37.93	39.05	37.70	38.38
10/28.11	35.57	33.52	35.47	33.89	34.61
	35.61	33.96	35.36	33.89	34.70
15/24.62	33.75	31.51	33.69	31.74	32.67
	33.75	31.59	33.52	31.76	32.65
20/22.11	32.50	30.43	32.20	30.10	31.23
	32.38	29.93	32.07	30.23	31.15
25/20.16	31.45	28.97	31.22	28.91	30.13
	31.32	28.63	30.97	29.02	29.98
50/14.15	28.19	25.64	27.81	24.97	26.65
	28.66	24.77	27.49	25.33	26.56
100/8.12	24.95	22.78	23.99	21.55	23.31
	24.94	21.85	23.70	21.48	22.99

Fig. 3. CPU run time for denoising of "Barbara" image (image size: 512×512)

Table 3. Denoising PSNR results in decibels of the algorithm proposed in [10] (OMP) and our algorithm (adaptive thresholding) in the presence of the uniform noise. Test image:"Barbara" image.

σ/PSNR	Our algorithm	OMP
2/42.11	43.57	43.61
5/34.16	37.87	37.94
10/28.13	33.91	34.01
15/24.61	31.65	31.69
20/22.11	30.01	30.04
25/20.18	28.77	28.73
50/14.15	25.03	24.86
100/8.13	21.62	21.90

our algorithm slightly outperforms the algorithm proposed in [10].

We compare the computational complexity of our algorithm and OMP algorithm using the CPU run time. Although CPU time is not an exact measure, it gives a rough estimation of complexity. The execution time of both algorithms for "Barbara" image of size 512×512 and different noise levels is presented in Fig. 3. All the experiments were done with an Intel Core 2 Duo processor 2.4GHz with 2GB of memory, using MATLAB R2006a environment and under Microsoft Windows Vista operating system. The CPU run time of our algorithm for an image of size 512×512 is about 14 seconds. As can be seen, the proposed algorithm is highly faster than OMP.

As was stated, the proposed algorithm is based on the assumption that the image is corrupted by zero mean white Gaussian noise. We conduct an experiment to investigate the denoising performance of the proposed algorithm if the noise is not Gaussian. Table 3 summarizes the denoising results when the image is corrupted by a zero mean uniform noise. The test image in this experiment is "Barbara" image. As can be seen, the proposed algorithm works fairly well, if the noise is uniform and not longer Gaussian. Another simulation is done to investigate the performance of our algorithm in presence of an exponential noise. Again the test image is "Barbara". The denoising results are listed in Table 4 in terms of PSNR. It is clear that denoising results of our algorithm are not satisfactory in this case. We can conclude that our algorithm has not promising denoising results if the noise is not Gaussian. It should be noted that denoising results of the OMP algorithm when the noise is exponential are also not satisfactory.

Figure 4-(a) and (b) shows a part of the noisy image of "Barbara" and the clean image which is denoised by our algorithm and Fig. 4-(c) shows the difference between the original and denoised images. In this experiment the variance of noise is 20. It can be seen that the proposed algorithm preserves the details of the image effectively while removing the noise.

Table 4. Denoising PSNR results in decibels of the algorithm proposed in [10] (OMP) and our algorithm (adaptive thresholding) in the presence of the exponential noise. Test image: "Barbara" image.

σ/PSNR	Our algorithm	OMP
2/39.11	39.76	39.77
5/31.14	32.58	32.57
10/25.12	27.07	27.05
15/21.61	23.79	23.75
20/19.09	21.41	21.36
25/17.16	19.56	19.51
50/11.14	13.77	13.71
100/5.13	7.90	7.91

5. CONCLUSION

In this paper we proposed an algorithm for image denoising based on redundant representations and adaptive thresholding for coefficient shrinkage. We mentioned that the proper threshold depends on statistical parameters of the noise, so we derived the statistical properties of the noise in the redundant transform domain. Simulation results showed that the proposed algorithm achieves the state-of-the-art denoising performance while it is computationally more efficient. This efficiency is mainly due to the fact that our algorithm needs only to calculate the minimum ℓ^2 norm representation under the redundant dictionary which can be done faster than finding the sparse representation of a signal.

6. REFERENCES

[1] R. G. Gonzalez and R. E. Woods, *Digital Image Processing*, Prentice Hall, 2007.

[2] D. L. Donoho, I. M. Johnstone, G. Kerkyacharian, and D. Picard, "Wavelet shrinkage: Asymptotia?," *J. Royal statist. Soc.*, vol. 51, no. 2, pp. 301–337, 1995.

[3] J. L. Starck, E. J. Candes, and D. L. Donoho, "The curvelet transform for image denoising," *IEEE Transactions on Image Processing*, vol. 11, no. 6, pp. 670–684, 2002.

[4] M. Lang, H. Guo, and J. E. Odegard, "Noise reduction using an undecimated discrete wavelet transform," *IEEE Signal processing Letters*, vol. 3, no. 1, pp. 10–12, 1996.

[5] R. Eslami and H. Radha, "The contourlet transform for image de-noising using cycle spinning," in *Asilomar Conference on Signals, Systems, and Computers*, 2003, pp. 1982–1986.

(a)   (b)

(c)
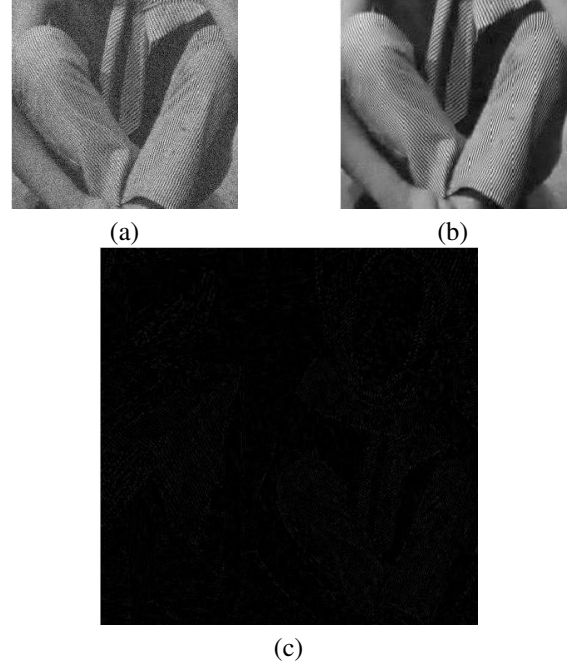
Fig. 4. (a) noisy Barbara image ($\sigma = 20$), (b) denoised image using the proposed algorithm, (c) difference between original and denoised images.

[6] M. Elad, "Why simple shrinkage is still relevant for redundant representations?," *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5559–5569, 2006.

[7] M. Elad, B. Matalon, and M. Zibulevsky, "Image denoising with shrinkage and redundant representations," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2006, pp. 17–22.

[8] D. L. Donoho, "For most large underdetermined systems of linear equations the minimal l^1-norm solution is also the sparsest solution," Tech. Rep., 2004.

[9] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 33–61, 1999.

[10] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Transactions on Image Processing*, vol. 15, no. 12, pp. 3736–3745, December 2006.

[11] K. Dabov, R. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising with block-matching and 3d filtering," in *Electronic Imaging06, Proc. SPIE 6064, no. 6064A-30*, 2006.