Image denoising using matched biorthogonal wavelets

Sanjeev Pragada and Jayanthi Sivaswamy
Center for Visual Information Technology,
International Institute of Information Technology, Hyderabad
{pragadask@research., jsivaswamy@}iiit.ac.in

Abstract

Current denoising techniques use the classical orthonormal wavelets for decomposition of an image corrupted with additive white Gaussian noise, upon which various thresholding strategies are built. The use of available biorthogonal wavelets in image denoising is less common because of their poor performance. In this paper, we present a method to design image-matched biorthogonal wavelet bases and report on their potential for denoising. We have conducted experiments on various image datasets namely Natural, Satellite and Medical with the designed wavelets using two existing thresholding strategies. Test results from comparing the performance of matched and fixed biorthogonal wavelets show an average improvement of 35% in MSE for low SNR values (0 to 18db) in every dataset. This improvement was also seen in the PSNR and visual comparisons. This points to the importance of matching when using wavelet-based denoising.

1. Introduction

The wavelet transform has been a powerful and widely used tool in image denoising because of its energy compaction and multiresolution properties. Denoising an image corrupted with additive white Gaussian noise was initially proposed in [4] by thresholding the wavelet coefficients. Subsequently, various decomposition strategies and thresholding schemes have been proposed [1], [8], [7]. However, most of these use classical orthogonal wavelets which are independent of the image and noise characteristics and focus on finding the best threshold.

Unlike the Fourier transform with its complex exponential basis, the wavelet transforms do not have a unique basis. Noting this point several attempts at designing matched wavelets have been made with the goal of match varying from match to a signal [2] and energy compaction [9] to maximizing the signal energy in the scaling sub-space [5]. The matched wavelets have been applied to feature extraction and compression problems. But as per our knowledge, no such work has been reported in particular for the image denoising problem.

Our work explores the utility of matched wavelets for denoising. In this paper, we report on a design methodology based on biorthogonal wavelets. We assume the noise to be additive white Gaussian and use the statistical properties of noise in the design. Specifically, we propose a design that will generate a biorthogonal wavelet bases for a given corrupted image which can be used for its denoising.

The paper is organized as follows: the design of adaptive biorthogonal wavelet bases is presented in section 2 followed by an algorithm for image denoising using matched wavelets and a brief review of thresholding strategies used, in section 3. In section 4 we present the obtained results and the corresponding discussion; we finally close with some conclusions and future work in section 5.

2. Image matched biorthogonal wavelets

We use the concept of *separable kernel* proposed by Mallat [6] in our design of matched wavelets for images. Hence, two sets of 1D matched wavelets are designed for two 1-D signals generated from the given image, obtained by row and column orderings instead of designing the two-dimensional matched wavelets. We have opted for separable kernel mainly for simplicity of the design procedure. Now the problem of finding image-matched biorthogonal wavelets is essentially one of designing 2-channel 1D FIR perfect reconstruction filter bank for each of the two 1-D signals satisfying some necessary and sufficient conditions. The design of FIR filters makes the obtained wavelets to be compactly supported.

Our criterion for matching is based on maximizing the projection of signal characteristics into the scaling subspace rather than the wavelet subspace, based on a knowledge of the noise characteristics. Such a matching criteria will lead to a higher SNR in the coarsest approximation subspace when a noisy signal is decomposed with the matched as opposed to a fixed wavelet. Thus a thresholding process which

978-0-7695-3476-3/08 $25.00 © 2008 IEEE
DOI 10.1109/ICVGIP.2008.95

passes the coarsest approximation sub-band and attenuates the rest of the sub-bands should decrease the amount of residual noise in the overall signal after the denoising process. A similar approach was proposed by Gupta *et al.* [5] for designing wavelets for signal and image compression. However, their wavelets would not work for the denoising applications as they are matched directly to the input image which is a noisy image in denoising applications i.e their technique would maximize the projection of noisy image features into the scaling subspace and not the clean image features. Next, we present our design in detail.

2.1. Design of two-channel 1D PR filter bank

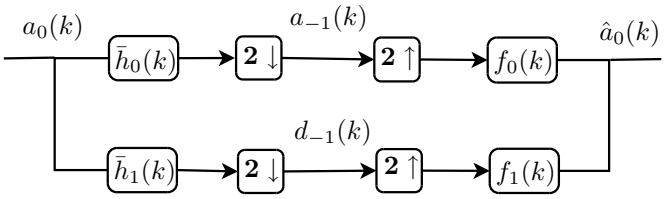


Figure 1. Two channel 1D maximally decimated PR filter bank:
bar on analysis filters denotes time reversal

Fig.1 shows a 1D two-channel maximally decimated filter bank. Given a noisy image we obtain two 1D noisy signals, one by row and the other by column ordering. Now we design a two channel 1D perfect reconstruction filter bank for each of these 1D noisy signals using our matching criteria. In our approach, the analysis high pass filter $h_1(k)$ is designed first and the bi-orthogonal relations are then used to design the remaining filters of the filter bank.

Design of analysis high pass filter

Given a 1D noisy signal, we design the analysis high pass filter using the matching criteria, that projects most of the pure signal energy from the noisy one into the scaling subspace and less into the wavelet subspace. The design is described below.

Consider the two channel filter bank shown in Fig.1. Let $a_0(k)$ be the discrete noisy signal of length L obtained from the noisy image and consider its values to be the coefficients of expansion in the scaling subspace $\mathbf{V_0}$. Now the signals $a_{-1}(k)$ and $d_{-1}(k)$ can be considered as the coefficients of expansion in the lower scaling subspace $\mathbf{V_{-1}}$ and wavelet subspace $\mathbf{W_{-1}}$ respectively. Let N be length of the filters. Hence we have,

$$a_{-1}(k) = \sum_{n=0}^{N-1} h_0(-n)a_0(2k-n) \tag{1}$$

$$d_{-1}(k) = \sum_{n=0}^{N-1} h_1(-n)a_0(2k-n) \tag{2}$$

Let ϕ(t) and ψ(t) be the scaling and wavelet functions respectively. The signal reconstructed from initial scaling space $\mathbf{V_0}$ is

$$z(t) = \sum_{k=0}^{L-1} a_0(k)\phi(t-k) \tag{3}$$

and the signal reconstructed using only the lower wavelet subspace $\mathbf{W\_1}$ be

$$\hat{z}(t) = \sum_{k=0}^{L/2-1} \frac{1}{\sqrt{2}} d_{-1}(k)\psi(t/2-k) \tag{4}$$

The error energy between the signals z(t) and $\hat{z}$(t) can be defined as

$$E = \int (z(t) - \hat{z}(t))^2 dt \tag{5}$$

Now maximizing the signal energy projected into the scaling subspace is equivalent to minimizing the projection into wavelet subspace and this in turn is equivalent to maximizing the error energy defined in Eq.5.

The basic set of equations required for the simplification of Eq.5 are given below.

$$\phi(t) = \sum_{n=0}^{N-1} \sqrt{2}h_0(n)\phi(2t-n) \tag{6}$$

$$\psi(t) = \sum_{n=0}^{N-1} \sqrt{2}h_1(n)\phi(2t-n) \tag{7}$$

$$\int \phi(t)\phi(t-p)dt = \delta(p) \tag{8}$$

$$\int \psi(t/2-p)\phi(t-m)dt = \sqrt{2}h_1(m-2p) \tag{9}$$

Substituting Eqs. 2, 3 and 4 into Eq. 5 and using the above four equations in its simplification will give rise to

$$E = \sum_{m=0}^{L-1} a_0(m)^2 - \sum_{=0p}^{N-1} \sum_{k=0}^{L/2-1} \sum_{=0\ q}^{N-1} h_1(p)h_1(q)a_0(2k+p)a_0(2k+q) \tag{10}$$

Now maximizing the simplified error energy w.r.t analysis high pass filter $h_1(k)$ [1] will lead to [5]

[1] The full derivation is too long to fit here. It can be provided if needed.

$$\sum_{p=0}^{N-1} h_1(p)\Big[\sum_{k=0}^{L/2-1} a_0(2k+p)a_0(2k+q)\Big] = 0 \qquad (11)$$

for r=0,1, . . . ,N-1

The filter designed using the Eq. 11 will project most of the input noisy signal energy into the scaling subspace but not the pure signal energy. Hence to match our criteria the equation 11 is modified accordingly as

$$\sum_{p=0}^{N-1} h_1(p)\Big[\sum_{k=0}^{L/2-1} x(2k+p)x(2k+r)\Big] = 0 \qquad (12)$$

where x(k) is assumed to be the pure signal and n(k) be AWGN with zero mean and σ^2 variance added to x(k) to get $a_0(k)$. Substituting x(k)=a(k)-n(k) in Eq.12 and using the i.i.d properties of n(k) we get

$$\sum_{p=0}^{N-1} h_1(p)\left[\Big[\sum_{k=0}^{L/2-1} a_0(2k+p)a_0(2k+r)\Big] - \Gamma\cdot\delta[p-r]\right] = 0 \qquad (13)$$

where Γ is given by

$$\Gamma = \frac{2}{L}\sum_{k=0}^{L/2-1} n^2(2k+r) \qquad (14)$$

The above equation can be linearly solved for $h_1(k)$ by estimating the autocorrelation of noise.

Estimation of noise statistics In the additive white Gaussian noise setting, the autocorrelation of noise is approximately equal to the noise variance σ^2. There are two scenarios to obtain this: either we have apriori information about the noise variance or it can be estimated from the given noisy image using a robust median estimator [4].

Design of the other three filters

The remaining three filters in Fig.1 are obtained from $h_1(n)$ using the biorthogonal relations and perfect reconstruction conditions [10], [3]. That is, compute $f_0(n)$ as

$$f_0(n) = (-1)^{n+1}h_1(M-n) \qquad (15)$$

where M is any odd delay.
Next, $h_0(n)$ can be calculated using two biorthogonal relations below

$$\sum_{n=0}^{N-1} h_0(n)f_0(n-2l) = \delta(l) \qquad \forall l \in Z \qquad (16)$$

$$\sum_{n=0}^{N-1} h_0(n)h_1(n) = 0 \qquad (17)$$

Additionally, we impose *p* vanishing moments on $f_1(n)$ as follows. These moments are

$$m_1(k) = \sum_{n=0}^{N-1} n^k f_1(n) = 0 \qquad fork = 0, 1, \ldots, p. \quad (18)$$

They can be transferred to $h_0(n)$ as

$$\sum_{n=0}^{N-1} (-n)^k h_0(n) = 0 \qquad fork = 0, 1, \ldots, p \qquad (19)$$

Equations 16, 17 and 19 can be solved simultaneously to get $h_0(n)$ with *p* vanishing moments.

Finally, $f_1(n)$ can be determined using

$$f_1(n) = (-1)^n h_0(M-n) \qquad (20)$$

Thus, all the four filters of the matched biorthogonal filter bank are derived as explained above.

3. Overall denoising algorithm

The overall filter design and denoising processes are summarized in the algorithm shown below.

Algorithm 1 Image Denoising using Matched Wavelets

1: Denote signals obtained by row and column ordering of noisy image as a_{0x} and a_{0y} respectively
2: Use a_{0x} as input to two channel filter bank shown in 1 and denote the filters thus designed as $h_{0x}, h_{1x}, f_{0x}, f_{1x}$
3: Now use a_{0y} as input and denote the filters thus designed as $h_{0y}, h_{1y}, f_{0y}, f_{1y}$
4: Obtain the 2D DWT of given noisy image using the 1D decomposition filters $h_{0x}, h_{1x}, h_{0y}, h_{1y}$ by separable kernel upto desired number of levels.
5: Threshold the detail subbands at each level using the chosen thresholding strategy and leave the coarsest approximation subband unthresholded.
6: Now obtain the reconstructed image from above coefficents using the 1D reconstruction filters $f_{0x}, f_{1x}, f_{0y}, f_{1y}$ by separable kernel.

The first level decomposition and reconstruction process using matched wavelets is shown in Fig.2. As the designed filters are one-dimensional we use the separable kernel [6] to compute the forward and reverse wavelet transforms. In the figure, LL are the first level approximation coefficients and LH, HL and HH are the first level detail coefficients. The first level approximation coefficients can be applied as

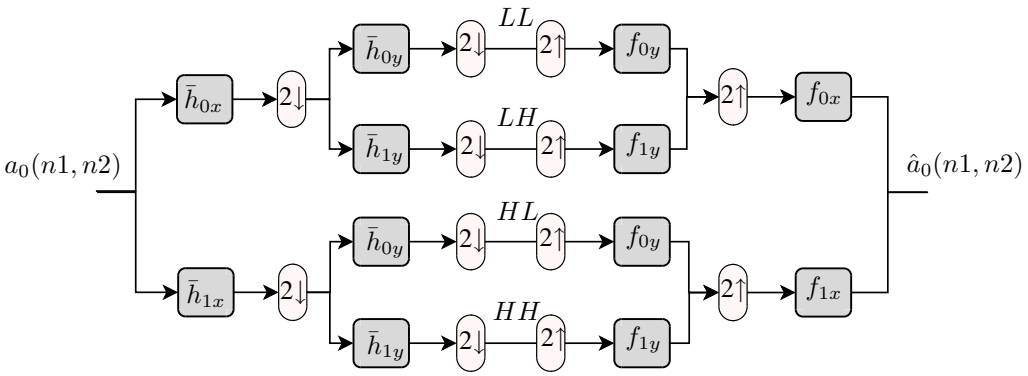


Figure 2. First level decomposition and reconstruction using separable kernel and matched filters; The bars on filters denote time reversal

the input to Fig.2 to get the second level coefficients and so on.

For the thresholding stage of the algorithm we have used two strategies namely *BayesShrink* proposed in [1] and *BiShrink* proposed in [8]. Since there are no thresholding strategies specially designed for biorthogonal wavelets, we have used the above strategies although they were originally designed for orthogonal wavelets. And since our main goal here is to compare the denoising performance of our matched biorthogonal wavelets with the fixed (CDF) biorthogonal wavelets, these strategies serve our purpose. We briefly describe them here for making our paper self-contained and clear.

BayesShrink - Let the observation model be $Y = X + V$, where Y, X and V are wavelet coefficients of noisy, original and noise images respectively with X and V independent of each other, hence

$$\sigma_Y^2 = \sigma_X^2 + \sigma^2$$

where σ_Y^2 is the variance of Y. Since Y is modeled as zero mean, σ_Y^2 can be found empirically by

$$\hat{\sigma}_Y^2 = \frac{1}{n^2} \sum_{i,j=1}^{n} Y_{ij}^2$$

where n×n is the size of the subband under consideration.

The variance of noise can be estimated using a robust median estimator presented in [4] as

$$\hat{\sigma} = \frac{Median(|Y_{ij}|)}{0.6745}, \qquad Y_{ij} \in subbandHH_1.$$

where HH_1 is the subband containing finest level diagonal details.

Now the threshold value is given by

$$\hat{T} = \frac{\hat{\sigma}^2}{\hat{\sigma}_X} \tag{21}$$

where

$$\hat{\sigma}_X = \sqrt{max(\hat{\sigma}_Y^2 - \hat{\sigma}^2, 0)}.$$

In the case that $\hat{\sigma}^2 \geq \hat{\sigma}_Y^2$, $\hat{\sigma}_X$ is taken to be 0. That is, $\hat{T}$ is ∞, or, in practice, $\hat{T} = max(|Y_{ij}|)$, and all coefficients are set to 0. This happens at times when σ is large.

To summarize, *BayesShrink* is the denoising strategy which performs soft-thresholding [4], with the data-driven, subband-dependent threshold given by Eq. 21.

BiShrink - Let w_2 represent the parent of w_1 (w_2 is the wavelet coefficient at the same spatial position as w_1, but at the next coarser scale). Then

$$y_1 = w_1 + n_1 \quad and \quad y_2 = w_2 + n_2$$

where y_1 and y_2 are noisy observations of w_1 and w_2, and n_1 and n_2 are noise samples. We can write

$$\mathbf{y} = \mathbf{w} + \mathbf{n}$$

where $\mathbf{w} = (w_1, w_2)$, $\mathbf{y} = (y_1, y_2)$ and $\mathbf{n} = (n_1, n_2)$.

The threshold value is given by

$$\hat{T} = \frac{\sqrt{3}\hat{\sigma}_n^2}{\hat{\sigma}_w} \tag{22}$$

where $\hat{\sigma}_n^2$ and $\hat{\sigma}_w^2$ are estimates of noise and signal variances respectively which can be obtained as in the *BayesShrink* strategy.

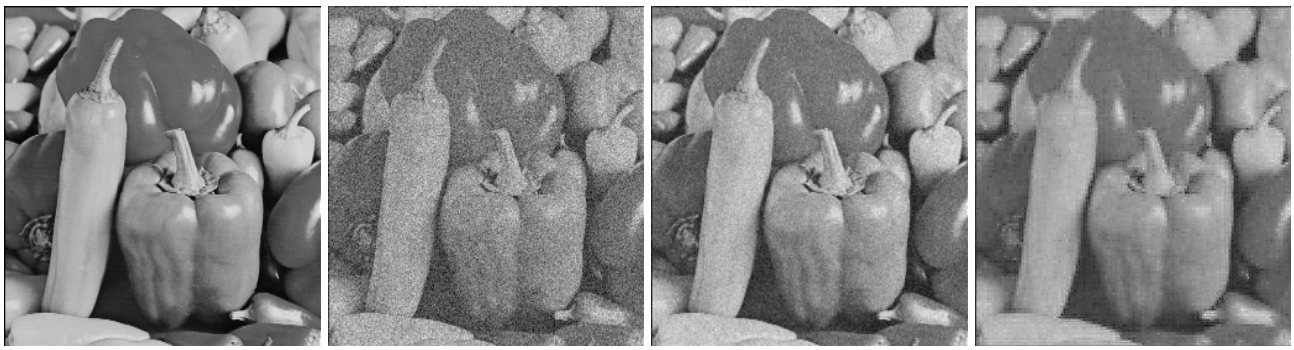


Figure 3. BayesShrink - From left to right: Original, Noisy peppers ($\sigma = 30$), Denoised using bior3.5, Denoised using matched wavelets

The bivariate shrinkage or thresholding function used in this strategy is

$$\hat{w}_1 = \frac{(\sqrt{y_1^2 + y_2^2} - \hat{T})+}{\sqrt{y_1^2 + y_2^2}}.y_1. \tag{23}$$

Here $(g)_+$ is defined as

$$(g)_+ = \begin{cases} 0 & \text{if } g < 0 \\ 0 & \text{otherwise} \end{cases}$$

To summarize, *BiShrink* is the denoising strategy which performs bivariate shrinkage given by Eq. 23, with the data-driven, subband-dependent threshold given by Eq. 22.

4. Experimental results and discussion

The denoising performance of our matched wavelet was tested on set of standard 8-bit grayscale images such as *Peppers*, *Elaine*, *Barbara*, *Baboon*, and *Goldhill* etc., and two other datasets namely Satellite and Medical(HRA) for various noise-levels. It is compared against existing fixed biorthogonal wavelets and orthogonal wavelets. We have used 4-level wavelet decomposition, and the coarsest approximation coefficients were *not* thresholded. We have tested the performance of our matched wavelets with two types of thresholding strategies: *BayesShrink* proposed in [1] and *BiShrink* proposed in [8] and described in the previous section.

In order to study the performance of matched wavelets w.r.t the frequency content of the image, the test images were also classified by their frequency distribution. For doing so, the PSD of the given image is calculated in three intervals of full spectrum such as i)0 to $\pi/3$, ii)$\pi/3$ to $2\pi/3$ and iii)$2\pi/3$ to π. Now the images were classified as low pass, band pass and high pass when PSD is maximum in interval1, interval2 and interval3 respectively. The experiments were carried on the set of low pass and high pass images only, because they are the ones which are found more often in general and in our datasets.

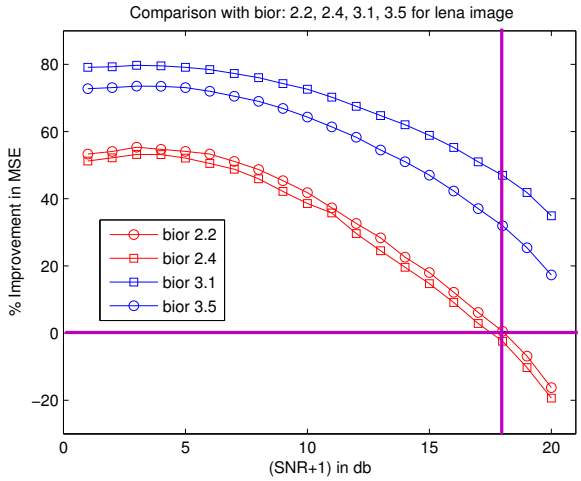


Figure 4. BayesShrink- Percentage improvement in MSE over fixed biorthogonal wavelets for image 'Lena'

4.1. Comparison against fixed biorthogonal wavelets

BayesShrink - The visual comparison of performances between matched and biorthogonal (bior3.5) wavelets is shown in Fig.3. This image is classified as low pass from the computed PSD values. The results show a visible improvement in the image quality in terms of noise content albeit with some smoothing. The PSNR comparisons against the fixed biorthogonal wavelets are shown in Table. 1. It shows the results for four natural images namely *Lena*, *Goldhill*, *Baboon* and *Barbara*. It can be seen from the table that for

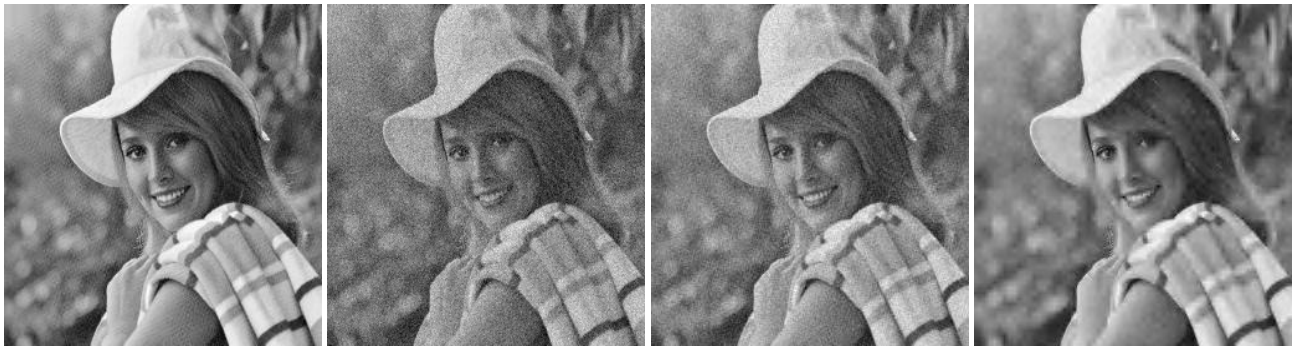


Figure 5. BiShrink - From left to right: Original, Noisy elaine ($\sigma = 20$), Denoised using bior3.1, Denoised using matched wavelets

higher values of noise standard deviation (>10) the matched wavelets are performing better than the fixed wavelets. On an average the PSNR improvement was found be approximately 2dB. The percentage improvement in MSE (α) obtained with matched wavelets is shown in Fig. 4 and is calculated as

$$\alpha = 100 * \frac{MSE_{fixed} - MSE_{matched}}{MSE_{fixed}}$$

BiShrink - Now we tested the performance of our matched wavelet with the *BiShrink* threshold. The visual results can be seen in Fig 5. The results are shown on a different image of the low pass set compared to image used for *BayesShrink* results. This is done to show the robustness of matched wavelets for various images as long as they are from the same frequency set. The percentage improvement in MSE with this threshold is shown in Fig 6. We observe that the improvement in MSE performance with *BiShrink* threshold is more or less same as with *BayesShrink* threshold. This indicates the robustness of matched wavelets w.r.t thresholding method. The PSNR comparisons against the fixed biorthogonal wavelets obtained with this method are shown in Table.2

The visual results for the satellite and medical images are shown in Fig 4.2. The results show the comparisons with the *bior2.4* and *bior2.2* wavelets for the noise standard deviations of 20 and 30 respectively. The thresholding technique used here is the *BiShrink*, as it was seen to outperform *BayesShrink* in most of the cases. Clearly it can be observed from the figure that matched wavelets are well suited for these kinds of application specific datasets also apart from the natural images.

From the computed percentage improvement values in Fig. 4 and Fig. 6 we can see that denoising with matched wavelets results in an improvement of 35% percent on average for lower SNR (0-18db) with this performance decreasing at higher SNR (>18db). The improvement in denoising

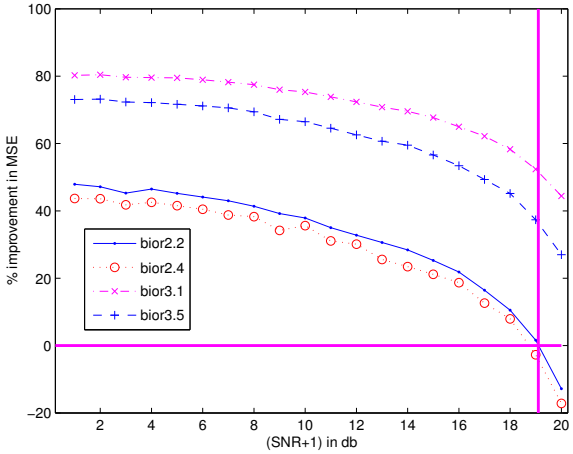


Figure 6. BiShrink- Percentage improvement in MSE over fixed biorthogonal wavelets for image 'Lena'

at low SNR cases is very much attractive. Hence we suggest this approach in lower SNR cases where there is the actual need of matching.

The decrease in the performance improvement at higher SNRs can be explained as follows. As the matched wavelets are adaptive to the noise conditions they perform equally well for full SNR range. But at higher SNR since the noise energy corrupting the image is low, the fixed wavelets perform better compared to their own performance at lower SNRs. And hence MSE improvement plot shows performance decrease as the SNR level increases. Anyway this behavior need not be of much concern because SNR (>18db) generally leads to a noise standard deviation (<10) which is very low and at such noise levels denoising itself is insignificant.

Table 1. BayesShrink - PSNR COMPARISON OF MATCHED WAVELET AND FIXED BIORTHOGONAL WAVELETS

σ	5	10	20	30	50	5	10	20	30	50
input PSNR	34.14	28.13	22.10	18.57	14.16	34.14	28.13	22.10	18.57	14.16
Wavelet	**Lena 512 X 512**					**Goldhill 512 X 512**				
Bior3.5	35.77	30.61	24.92	21.47	17.10	35.22	30.27	24.74	21.35	17.03
Bior2.2	36.22	31.65	26.52	23.36	19.32	35.38	31.02	26.28	23.19	19.15
Bior2.6	36.26	31.75	26.69	23.54	19.31	35.46	31.08	26.31	23.34	19.24
Matched	31.15	29.69	26.74	24.57	21.73	30.75	28.46	26.67	24.13	21.54
Wavelet	**Baboon 512 X 512**					**Barbara 512 X 512**				
Bior3.5	32.66	28.36	23.87	20.87	16.77	34.81	29.47	24.36	21.14	16.92
Bior2.2	32.57	28.47	24.46	21.85	18.39	35.19	29.93	25.30	22.46	18.72
Bior2.6	32.83	28.62	24.55	21.92	18.47	35.28	30.01	25.43	22.53	18.84
Matched	27.43	26.01	24.72	22.83	20.99	30.34	27.62	25.81	23.50	20.02

Table 2. BiShrink - PSNR COMPARISON OF MATCHED WAVELET AND FIXED BIORTHOGONAL WAVELETS

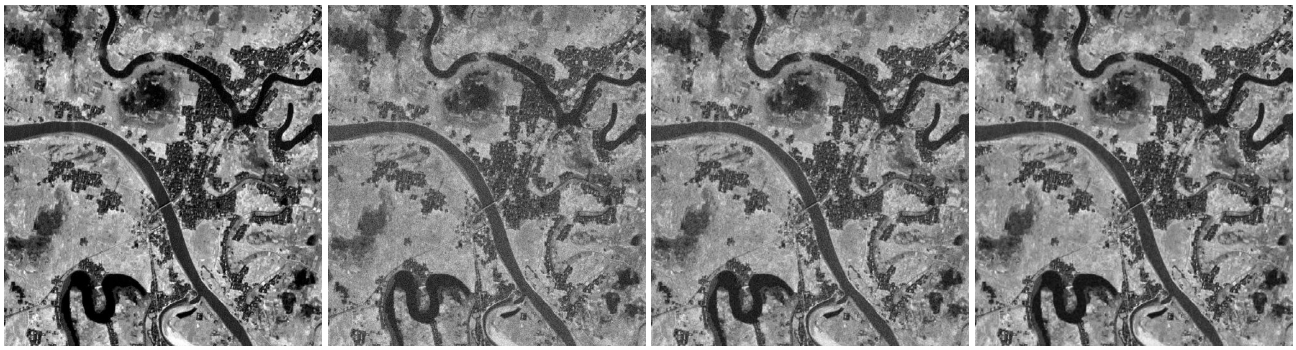
σ	5	10	20	30	50	5	10	20	30	50
input PSNR	34.14	28.13	22.10	18.57	14.16	34.14	28.13	22.10	18.57	14.16
Wavelet	**Lena 512 X 512**					**Goldhill 512 X 512**				
Bior3.5	36.23	31.12	25.46	21.89	17.56	35.72	30.74	25.13	21.77	17.48
Bior2.2	36.71	32.16	27.09	23.80	19.73	35.83	31.52	26.74	23.61	19.59
Bior2.6	36.74	32.25	27.21	23.94	19.76	35.94	31.59	26.80	23.70	19.62
Matched	31.85	31.09	27.34	25.07	22.23	31.15	29.16	27.27	25.43	22.94
Wavelet	**Baboon 512 X 512**					**Barbara 512 X 512**				
Bior3.5	33.16	28.86	24.31	21.24	17.16	35.33	29.97	24.89	21.56	17.32
Bior2.2	33.07	28.98	24.96	22.25	18.79	35.69	30.43	25.80	22.87	19.13
Bior2.6	33.33	29.12	25.06	22.33	18.87	35.78	30.52	25.94	22.95	19.28
Matched	29.03	26.66	25.42	23.34	21.49	30.94	28.22	26.43	24.01	21.54

4.2. Comparison against fixed orthogonal wavelets

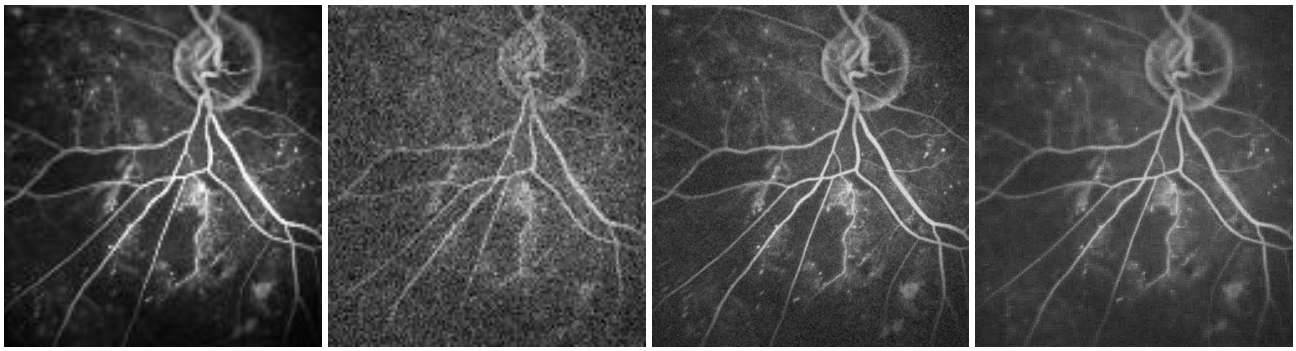
The performance of proposed biorthogonal wavelets was also compared with fixed orthogonal wavelets just for completeness. It was found that there was no improvement in denoising performance in this case. The reasons for this can be explained as follows. Firstly, in all our comparisons we used the thresholding techniques proposed in [1] and [8] which were originally designed for orthogonal wavelet decomposition and hence they can be expected to be sub-optimal for denoising with biorthogonal wavelets. However, we chose them as they are sub-band adaptive and in order to test the merit of matching in general. Notwithstanding this fact, the better performance of the matched wavelet against fixed biorthogonal and not orthogonal wavelet suggests that improved denoising is possible even with biorthogonal wavelet provided it is matched to the clean image. Secondly, the noise distribution across sub-bands is taken to be uniform (as per orthogonal decomposition) in the threshold derivation in [1]. Our investigations showed this assumption to be not true with biorthogonal decomposition.

5. Conclusions

In this paper, we have explored the utility of using image-matched biorthogonal filters for denoising. Our design of image-matched biorthogonal wavelet bases uses the constraint that most of the energy of clean image is projected into scaling subspace rather than the wavelet subspace. We have compared denoising performance of our matched wavelets with CDF biorthogonal wavelets with two well-known thresholding strategies for various image datasets. The results show that adapted biorthogonal



(a) Satellite: Original (b) Satellite: Noisy($\sigma = 20$) (c) Satellite: Denoised_Bior2.4 (d) Satellite: Denoised_Matched



(e) Medical: Original (f) Medical: Noisy($\sigma = 30$) (g) Medical: Denoised_Bior2.2 (h) Denoised_Matched

Figure 7. Visual Comparisons for Satellite and Medical Images

wavelets performed much better denoising than the available biorthogonal wavelets for low SNR i.e where the actual need for adaptation arises. Also the results show that these matched wavelets can be used for denoising of variety of image datasets such as Satellite and Medical and also show their robustness w.r.t thresholding technique. Developing a thresholding technique specifically suited for biorthogonal wavelets needs investigation as it may improve the denoising of matched wavelets against fixed orthogonal wavelets also. The design of matched orthogonal wavelets is another line that is presently under investigation.

References

[1] G. Chang, B. Yu, and M. Vetterli. Adaptive wavelet thresholding for image denoising and compression. *IEEE Transactions on Image Processing*, 9:1532–1546, September 2000.

[2] J. O. Chapa and R. M. Rao. Algorithms for designing wavelets to match a specified signal. *IEEE Transactions on Signal Processing*, 2000.

[3] I. Daubechies. *Ten Lectures on Wavelets*. Philadelphia, PA: SIAM, 1992.

[4] D. L. Donoho and I. M. Johnstone. Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81:425–455, March 1994.

[5] A. Gupta, S. D. Joshi, and S. Prasad. A new approach for estimation of statistically matched wavelet. *IEEE Transactions on Signal Processing*, 53:1778–1793, May 2005.

[6] S. Mallat. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 11:674–693, July 1989.

[7] J. Portilla, V. Strela, M. Wainwright, and E. Simoncelli. Image denoising using scale mixtures of gaussians in the wavelet domain. *IEEE Transactions on Image Processing*, 12:1338–1351, November 2003.

[8] L. Sendur and I. W. Selesnick. Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency. *IEEE Transactions on Signal Processing*, 50:2744–2756, November 2002.

[9] A. H. Twefik, D. Sinha, and P. Jorgensen. On the optimal choice of a wavelet for signal representation. *IEEE Transactions on Information Theory*, 38:747–765, February 1992.

[10] M. Vetterli and J. Kovacevic. *Wavelets and Subband Coding*. Eaglewood Cliffs, NJ: Prentice Hall, 1995.