

Design of Image Adaptive Wavelets for Denoising Applications

Sanjeev Pragada and Jayanthi Sivaswamy
Center for Visual Information Technology

International Institute of Information Technology - Hyderabad, Hyderabad 500032, India

Email: {pragadask@research., jayanthi@}iiit.ac.in

Abstract—The traditional wavelet based denoising techniques use wavelets which are independent of the image under consideration and the level of noise present in the image. As these techniques which are based on thresholding of wavelet coefficients, rely mainly on the energy compaction and the sparseness of the wavelet representation used, the choice of wavelet is an important consideration. In this paper, for the first time we have proposed a method for designing biorthogonal wavelets for denoising applications from the image at hand using a separable kernel. The denoising performance of this image adapted wavelet is compared primarily against the available biorthogonal wavelets in terms of MSE, PSNR and visual quality, using *BayesShrink* method which is the current state-of-the-art technique in the class of pointwise approaches for image denoising. We have achieved a average improvement of 35% in MSE, over the available biorthogonal wavelets for the low input SNR values (0 to 18db); that is where the actual need for adaptation arises. The results show the efficacy of the approach and the need for adaptive wavelet bases in image denoising.

I. INTRODUCTION

Digital images are often degraded by noise in their acquisition and/or transmission phase. The goal of a image denoising algorithm is to recover the clean image from its noisy version by removing noise and retaining as much as possible the image information. The wavelet transform because of its multiresolution and sparse representations, has been shown to be a powerful tool to achieve these goals. The motivation is that small wavelet coefficients are more likely due to noise and the large coefficients due to the image features.

The interest in wavelet based denoising evolved from the seminal work on signal denoising via *wavelet thresholding or shrinkage* of Donoho and Johnstone in the additive white gaussian noise setting [1], [2]. These authors have proposed two thresholding strategies in which each wavelet coefficient is compared against a threshold; if the coefficient is smaller than the threshold, it is set to zero, otherwise it is kept (*hard thresholding*) or modified (*soft thresholding*) [3]. Initially, they have given a mechanism for finding a universal threshold value known as *VisuShrink* and showed it to be near-optimal in the minimax sense [1].

However, for denoising images, *Visushrink* is found to yield an overly smoothed estimate, as it is derived under the constraint that with high probability the estimate should be at least as smooth as the image. To overcome this problem, the same authors have proposed a new method called *SUREShrink* which uses a hybrid of the universal threshold and the SURE

threshold, derived from minimizing Stein's unbiased risk estimate [4], and has been shown to perform well. Later on, there has been a fair amount of research on threshold value selection for image denoising. One of the most popular was proposed by Chang *et al.*, who derived their threshold in a Bayesian framework, assuming a generalized Gaussian distribution for the wavelet coefficients and is called *BayesShrink* [5]. This method has a better MSE performance than *SUREShrink*.

In the more recent literature, people have considered the intra and interscale dependencies among the wavelet coefficients and have shown to achieve better denoising performance compared to the pointwise approach. In addition, increasing the redundancy of the wavelet transform has substantially improved the denoising performance. In these views the three main techniques are the *BiShrink* proposed by Sendur *et al.* [6], *ProbShrink* proposed by Pizurica *et al.* [7] and *BLS-GSM* proposed by Portilla *et al.* [8]. However, inspite of the improved performances with the use of scale dependencies and redundancy of the wavelet transform, we have used the point-wise techniques and non-redundant wavelet transforms for our comparison because of their simplicity and speed.

Our work aims at designing a image matched wavelet which can be used in denoising applications. The motivation for this is that, unlike the Fourier methods, wavelet transforms do not have a unique basis and hence it is natural to seek a wavelet that is best in a particular context. We have explored this fact for the image denoising problem. Over the last decade, there has been a fair amount of research on designing matched wavelets for a particular context [9], [10], [11]. But no work has been reported for the image denoising application in particular. The work done so far in the wavelet based denoising has targeted towards finding the best threshold but not towards finding the best wavelet. We, in this paper have concentrated on the latter problem and our results show the need for adaptive wavelets in image denoising.

In image denoising application, it is desirable that the wavelet adapts to the image and not to the noise. We present a technique to design wavelet from noisy image which adapts only to the clean image using statistical properties of noise. These properties help in designing a wavelet which can effectively deal with the noise present in the image, by improving the differential information between the image and noise in the wavelet domain. We have compared denoising performance of our designed biorthogonal wavelets with that of existing set

of biorthogonal wavelets using *BayesShrink* method which is the current state-of-the-art technique in the class of pointwise approaches.

This paper is organized as follows. In Section II, we present our basic work of designing bi-orthogonal wavelets for image denoising applications. In Section III, the overall process of image denoising using our matched wavelets is presented. Experimental results on several images and discussions are given in Section IV and finally Section V deals with the conclusions and future work.

II. DESIGN OF IMAGE MATCHED WAVELETS

In this section we will describe the algorithm for designing matched wavelets for image denoising. The image matched wavelets are designed using the *separable kernel*. i.e two sets of 1D matched wavelets are designed, one for row ordering of the image, say a_{0x} (arranging all rows side by side) which contains variation in the horizontal direction only and the other for column ordering of the image, say a_{0y} (arranging all columns one below the other and then taking its transpose) which contains variation in the vertical direction only. Now the design technique for finding 1D matched wavelets for denoising is presented. We use this technique two times, once with a_{0x} as input and obtain the four corresponding filters ($h_{0x}, h_{1x}, f_{0x}, f_{1x}$) and next with a_{0y} as input and obtain the four corresponding filters ($h_{0y}, h_{1y}, f_{0y}, f_{1y}$).

Our criterion for matching is based on maximizing the projection of clean image characteristics in to the scaling subspace rather than the wavelet subspace, using the noise characteristics. Using this matching criteria, we can guarantee that the noise in the coarsest approximation subspace will be less. And, as the thresholding techniques usually leave the coarsest approximation without thresholding, it means that our matching criteria is decreasing the amount of residual noise in the denoising process. A similar approach was proposed by Gupta *et al.* [11] and their designed wavelets have shown to be having profound application in 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 the 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. Now we present various stages of our algorithm.

A. Design of 1D Matched Biorthogonal Bases for image denoising

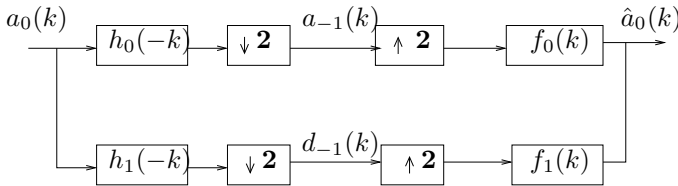


Fig. 1. Two channel 1D maximally decimated PR filter bank

Fig.1 shows a 1D two-channel maximally decimated filter bank. We use this filter bank for designing 1D matched

wavelets using a_{0x} and a_{0y} and then by using separable kernel we obtain the image matched wavelets which are suitable for denoising.

We first design the analysis highpass filters in each case and use the bi-orthogonal relations to design the remaining filters. Here, we shall talk in terms of 1D signals as we are designing 1D matched wavelets.

Design of Analysis High Pass Filter: Using the estimated noise variance, we shall design this filter such that our matching criterion is satisfied. Consider the scenario, where the signal is corrupted with white gaussian noise and we want to remove noise using signal adapted wavelet technique. Let $x(t)$ be the pure signal and $n(t)$ be WGN with independent and identical distribution (i.i.d) with zero mean and σ^2 variance. The noisy signal $a(t)$ is given by,

$$a(t) = x(t) + n(t) \quad (1)$$

The sampled value of the signal $a(t)$ is applied as a input to the 2-channel filter bank structure shown in Fig.1

$$a_0(k) = x(k) + n(k) \quad (2)$$

Let $\hat{a}(t)$ be the signal reconstructed using only the detail coefficients $d_{-1}(k)$ and N be the length of the filters in Fig 1. Now the analysis filter which maximizes the error energy between $a(t)$ and $\hat{a}(t)$ is obtained using the following equation which is given by Gupta *et al.* [11].

$$\sum_{p=0}^{N-1} h_1(p) \left[\sum_{k=0}^{L/2-1} a_0(2k+p)a_0(2k+r) \right] = 0 \quad (3)$$

For $r=0,1,\dots,j-1,j+1,\dots,N-1$.

where j is the index of middle coefficient of $h_1(k)$.

The wavelet designed using the above equation is matched to both signal and noise, where as our goal is to design the filter which is matched only to the signal. To achieve our goal the equation (3) is modified as below

$$\sum_{p=0}^{N-1} h_1(p) \left[\sum_{k=0}^{L/2-1} x(2k+p)x(2k+r) \right] = 0 \quad (4)$$

For $r=0,1,\dots,j-1,j+1,\dots,N-1$.

where j is the index of middle coefficient of $h_1(k)$.

Solving the above equation for h_1 directly, will need the knowledge of pure signal which we don't know. So we need to modify equation (4) further so that it can be solved directly.

Now, consider the bracketed term in the equation (3)

$$f(p,r) = \frac{2}{L} \sum_{k=0}^{L/2-1} a_0(2k+p)a_0(2k+r) \quad (5)$$

For $p=0,1,\dots,N-1$ and $r=0,1,\dots,N-1$

Using equation (2) we can write,

$$f(p,r) = \frac{2}{L} \sum_{k=0}^{L/2-1} [x(2k+p)+n(2k+p)][x(2k+r)+n(2k+r)] \quad (6)$$

$$\begin{aligned}
&= \frac{2}{L} \sum_{k=0}^{L/2-1} x(2k+p)x(2k+r) + \frac{2}{L} \sum_{k=0}^{L/2-1} x(2k+p)n(2k+r) \\
&+ \frac{2}{L} \sum_{k=0}^{L/2-1} n(2k+p)x(2k+r) + \frac{2}{L} \sum_{k=0}^{L/2-1} n(2k+p)n(2k+r)
\end{aligned}$$

Since we assumed that noise is i.i.d and uncorrelated to signal we have,

$$\frac{2}{L} \sum_{k=0}^{L/2-1} [x(2k+p)n(2k+r)] = 0 \quad (7)$$

$$\frac{2}{L} \sum_{k=0}^{L/2-1} [n(2k+p)x(2k+r)] = 0 \quad (8)$$

Putting these values in the equation (6), we have

$$f(p, r) = \frac{2}{L} \sum_{k=0}^{L/2-1} x(2k+p)x(2k+r) + \frac{2}{L} \sum_{k=0}^{L/2-1} n(2k+p)n(2k+r) \quad (9)$$

Now as auto-correlation function of white Gaussian noise is zero for any non-zero lag,

$$\begin{aligned}
\frac{2}{L} \sum_{k=0}^{L/2-1} n(2k+p)n(2k+r) &= \frac{2}{L} \sum_{k=0}^{L/2-1} n^2(2k+r) \quad ; \text{if } p = r \\
&= 0 \quad ; \text{otherwise} \quad (10)
\end{aligned}$$

Let

$$\Gamma = \frac{2}{L} \sum_{k=0}^{L/2-1} n^2(2k+r) \quad (11)$$

Substituting equation (10) and equation (11) in equation (9), we have

$$f(p, r) = \frac{2}{L} \sum_{k=0}^{L/2-1} x(2k+p)x(2k+r) + \Gamma \cdot \delta[p-r] \quad (12)$$

Thus

$$\frac{2}{L} \sum_{k=0}^{L/2-1} x(2k+p)x(2k+r) = f(p, r) - \Gamma \cdot \delta[p-r] \quad (13)$$

Substituting above equation and equation (5) in to equation (4), we have

$$\sum_{p=0}^{N-1} h_1(p) \left[\left[\sum_{k=0}^{L/2-1} a_0(2k+p)a_0(2k+r) \right] - \Gamma \cdot \delta[p-r] \right] = 0 \quad (14)$$

For $r=0,1,\dots,j-1,j+1,\dots,N-1$.

where j is the index of middle coefficient of $h_1(k)$.

where Γ is given by equation (11) and can be calculated by assuming a white Gaussian noise with zero mean and known variance.

Design of Remaining Filters in the Matched Biorthogonal Filter Bank: Now the remaining three filters in Fig.1 are obtained using $h_1(n)$ and the biorthogonal relations. That is, Compute $f_0(n)$ using

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

where M is any odd delay.

Now $h_0(n)$ can be calculated using the following two perfect reconstruction conditions and the vanishing moments imposed on $f_1(n)$

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

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

Now let's say we have imposed p vanishing moments on $f_1(n)$ then,

$$m_1(k) = \sum_{n=0}^{N-1} n^k f_1(n) = 0 \quad \text{for } k = 0, 1, \dots, p \quad (18)$$

which is transferred to $h_0(n)$ as

$$\sum_{n=0}^{N-1} (-n)^k h_0(n) = 0 \quad \text{for } k = 0, 1, \dots, p \quad (19)$$

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

Now, $f_1(n)$ can be easily found out from equation

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

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

III. EXPERIMENTAL RESULTS AND DISCUSSION

Now we use the standard wavelet based image denoising algorithm to compare the performance of our matched wavelets with that of fixed biorthogonal wavelets. It includes calculation of forward wavelet transform for the input noisy image and later thresholding of these coefficients with a properly designed threshold and then calculation of reverse wavelet transform of the thresholded coefficients to get the denoised image. The forward and reverse wavelet transforms can be calculated using the eight one dimensional filters ($h_{0x}, h_{1x}, f_{0x}, f_{1x}$) and ($h_{0y}, h_{1y}, f_{0y}, f_{1y}$) which can be designed using the method described in previous section and the non-redundant separable wavelet system shown in Fig 2. The thresholding strategy used for our comparison is the *BayesShrink* proposed by Chang *et al* [5].

The denoising performance of our matched wavelet is compared against existing biorthogonal wavelets for a representative set of standard 8-bit grayscale images such as *lena*, *barbara*, *baboon*, and *goldhill*, for various noise-levels. In our simulations, we have used 4 level wavelet decomposition obtained by decomposing the LL subband of Fig 2 further and

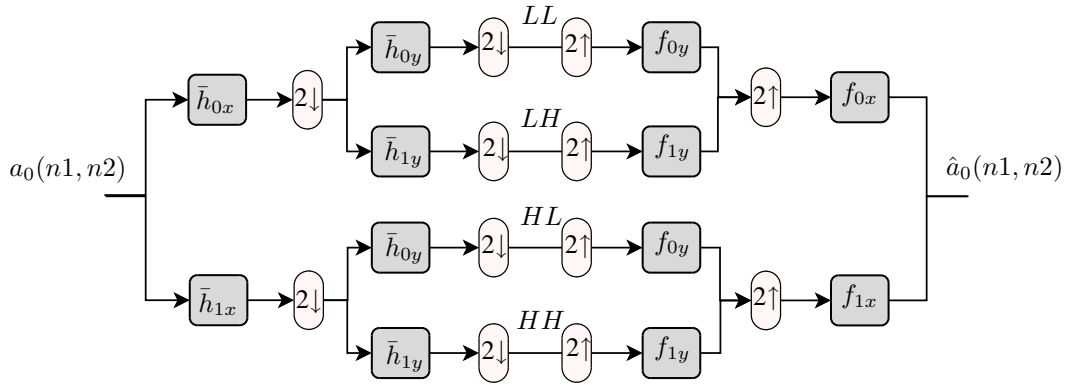


Fig. 2. First level decomposition and reconstruction using separable kernel and matched filters; The bars on filters denote time reversal



Fig. 3. From left to right: Original Lena , Noisy Lena ($\sigma = 30$) , Denoised using bior3.5, Denoised using matched wavelets

the coarsest approximation coefficients are left unthresholded. Although the thresholding technique, we have used in our simulations [5] was originally designed for orthonormal wavelets, selection of such a threshold is justified as our main motive is to compare the performances of matched wavelets versus fixed biorthogonal wavelets. The results of visual comparison for the lena image are shown in Fig 3 , MSE comparisons for the lena image are shown in Figs 4 and 5, and PSNR comparisons for the representative set are tabulated in Table I.

As we can see from Figs 4 and 5, proposed technique performs on an average 35% percent better denoising as compared to non-adaptive biorthogonal wavelets at lower SNR (0-18db) and at higher SNR (>18db), percentage improvement in MSE with respect to non-adaptive wavelet transform decreases. This is a very good improvement in the low SNR cases and that is where the idea of matching is very much required. This is also evident from PSNR comparisons in Table I.

At higher SNR, the energy of the noise is very less compared to that of signal. So the wavelet coefficients due to noise are very small and are spread out along all scales as the noise assumed is random, where as coefficients due to signal are very high and most of them exist only for few scales. In this way it is very easy to distinguish signal and noise in wavelet space

by using thresholding technique. Thus, the use of matched biorthogonal wavelets did not yield superior performance to non-adaptive biorthogonal wavelets in this range.

However, at lower SNR, the energy of the noise is comparable to the signal. Here, if the correlation between the used wavelet and the image is more, the wavelet representation of the image will be more compact and the distinguishing information between the wavelet coefficients of the signal and noise will be more strong which will help to remove noise in the wavelet space. Thus in this range our adaptive biorthogonal wavelets are performing better than the non-adaptive ones.

IV. CONCLUSION

In this paper, we have exploited the fact that wavelet bases are not unique, in designing the best biorthogonal bases for image denoising. We have designed image adaptive biorthogonal wavelet bases from the constraint that most of the energy of clean image should go in to scaling subspace rather than the wavelet subspace. The results showed that adapted biorthogonal wavelets performed much better denoising than the available biorthogonal wavelets for the low input SNR values. However the performance improvement decreased for the high SNR values as explained above. We suggest this approach when noise energy dominates signal energy (low

TABLE I
PSNR COMPARISON OF MATCHED WAVELET AND FIXED BIORTHOGONAL WAVELETS

σ	5	10	15	20	25	30	50	100	5	10	15	20	25	30	50	100
input PSNR	34.14	28.13	24.60	22.10	20.16	18.57	14.16	8.12	34.14	28.13	24.60	22.10	20.16	18.57	14.16	8.12
Wavelet	Lena 512 X 512								Goldhill 512 X 512							
Bior3.5	35.77	30.61	27.31	24.92	23.03	21.47	17.10	11.09	35.22	30.27	27.09	24.74	22.94	21.35	17.03	11.13
Bior2.2	36.22	31.65	28.81	26.52	24.92	23.36	19.32	13.34	35.38	31.02	28.27	26.28	24.61	23.19	19.15	13.38
Bior2.6	36.26	31.75	28.87	26.69	25.07	23.54	19.31	13.63	35.46	31.08	28.42	26.31	24.71	23.34	19.24	13.57
Matched	31.15	29.69	28.25	26.74	25.57	24.57	21.73	16.80	30.75	28.46	27.05	26.67	25.01	24.13	21.54	16.68
Wavelet	Baboon 512 X 512								Barbara 512 X 512							
Bior3.5	32.66	28.36	25.82	23.87	22.22	20.87	16.77	11.01	34.81	29.47	26.53	24.36	22.59	21.14	16.92	11.02
Bior2.2	32.57	28.47	26.18	24.46	23.05	21.85	18.39	13.12	35.19	29.93	27.20	25.30	23.80	22.46	18.72	13.09
Bior2.6	32.83	28.62	26.25	24.55	23.10	21.92	18.47	13.10	35.28	30.01	27.33	25.43	23.95	22.53	18.84	13.28
Matched	27.43	26.01	25.68	24.72	24.21	22.83	20.99	15.15	30.34	27.62	26.02	25.81	24.2	23.5	20.02	16.15

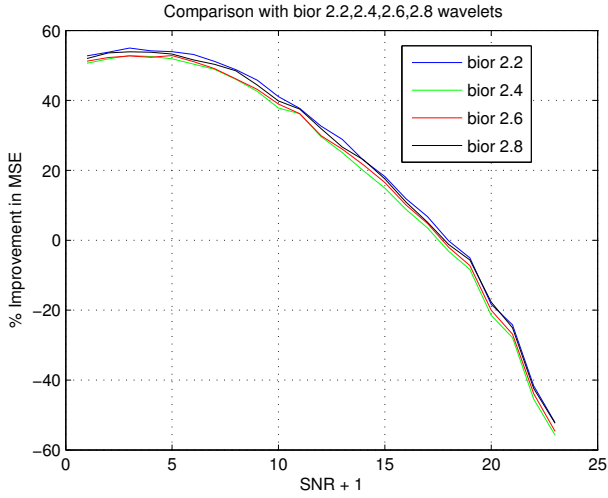


Fig. 4. Showing % improvement in MSE over bior2.* series wavelets for Lena image

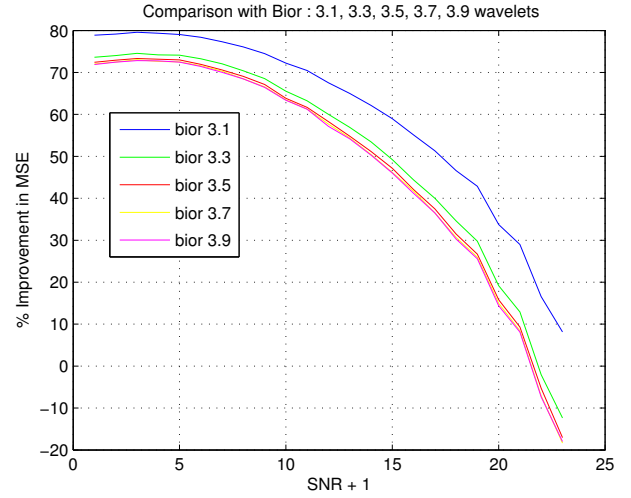


Fig. 5. Showing % improvement in MSE over bior3.* series wavelets for Lena image

SNR), because that is where the actual need for adaptation arises. Although our focus in this paper is on the adaptive biorthogonal wavelets, the adaptation can also be tried for orthonormal wavelets for yielding better denoising performance than the existing orthonormal wavelets.

REFERENCES

- [1] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994.
- [2] —, "Adapting to unknown smoothness via wavelet shrinkage," *Journal of the American Statistical Association*, vol. 90, no. 432, pp. 1200–1224, July 1995.
- [3] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 613–627, 1995.
- [4] C. M. Stein, "Estimation of the mean of a multivariate normal distribution," *The Annals of Statistics*, vol. 9, no. 6, pp. 1135–1151, 1981.
- [5] G. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Transactions on Image Processing*, vol. 9, no. 9, pp. 1532–1546, 2000.
- [6] L. Sendur and I. W. Selesnick, "Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency," *IEEE Transactions on Signal Processing*, vol. 50, no. 11, pp. 2744–2756, 2002.
- [7] A. Pizurica and W. Philips, "Estimating the probability of the presence of a signal of interest in multiresolution single- and multiband image denoising," *IEEE Transactions on Image Processing*, vol. 15, no. 3, pp. 654–665, 2006.
- [8] 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*, vol. 12, no. 11, pp. 1338–1351, 2003.
- [9] A. H. Tewfik, D. Sinha, and P. Jorgensen, "On the optimal choice of a wavelet for signal representation," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 747–765, 1992.
- [10] J. O. Chapa and R. M. Rao, "Algorithms for designing wavelets to match a specified signal," *IEEE Transactions on Signal Processing*, vol. 48, no. 12, pp. 3395–3406, 2000.
- [11] A. Gupta, S. D. Joshi, and S. Prasad, "A new approach for estimation of statistically matched wavelet," *IEEE Transactions on Signal Processing*, vol. 53, no. 5, pp. 1778–1793, 2005.