Denoising Via Block Wiener Filtering in Wavelet Domain

Vasily Strela

Abstract. In this paper we describe a new method for image denoising. We analyze statistical properties of the wavelet coefficients of natural images. It turns out that there is a strong local covariance structure introduced by the edges. We suggest a model for this covariance which allows us to estimate it from the noisy image. Then Wiener filter is employed in order to remove the noise.

We compare our approach to other noise removal techniques. Wiener-wavelet denoising produces superior results both visually and in terms of mean square error.

1. Introduction

A good model of the signal statistics is essential in many applications. This paper describes a simple and effective model for the covariance structure of natural images. We use this model for noise removal.

There are two powerful techniques to reduce the noise level in a signal: Wiener filtering [3] and wavelet thresholding [2]. Wiener filtering is a linear procedure. Wavelet thresholding is nonlinear. Classical versions of both methods tend to blur edges in images. We try to blend these two approaches in order to improve the performance. Our idea is to apply Wiener filter to blocks of wavelet coefficients. This requires an estimate of the covariance matrix of each block. We obtain it adaptively which makes the whole procedure nonlinear (similar to thresholding).

Many authors argue that the distribution of wavelet coefficients of images is strongly non-Gaussian. In particular the histogram has a sharp peak at zero. Common approach is to model the distribution as a generalized Gaussian $e^{-|y/\lambda|^p}$ [1, 4, 6, 7, 8]. Instead of modeling statistics of all wavelet coefficients together we suggest that each subband has slightly different statistics. Moreover, it is important to concentrate on blocks of wavelet coefficients.

Why do we choose to work with blocks? It is know that the wavelet transform acts as an edge detector. In images edges represent features. Each feature corresponds to a block of wavelet coefficients. Working with these blocks is natural in order to preserve the edges (features) better.

It can be proved that if both the signal and the noise are Gaussian then Winer filtering is the best possible mean square estimator. Our approach is to assume that in each subband blocks of wavelet coefficients are Gaussian vectors with slowly changing covariance matrix. We concentrate on creating a model for the covariance of the blocks and hope that Wiener filtering will be close to optimal. Experimental results presented in section 5 confirm our point of view.

2. Wiener Filter

Suppose a vector $\mathbf{S}$ is corrupted by Gaussian white noise with variance σ^2 and mean 0, $\mathbf{X} = \mathbf{S} + \sigma\mathbf{Z}$. Wiener filtering is the following linear procedure:

$$\widehat{\mathbf{X}} = \sum_m \frac{\beta_m^2}{\beta_m^2 + \sigma^2} \langle \mathbf{X}, \mathbf{g}_m \rangle \mathbf{g}_m \, . \tag{1}$$

Here β_m and $\mathbf{g}_m$ are eigenvalues and eigenvectors of the covariance matrix (Karhuen-Loeve transform) of $\mathbf{S}$. If $\mathbf{S}$ is Gaussian then $\widehat{\mathbf{X}}$ is the best mean square estimate of $\mathbf{S}$ (see for example [3]). In order to apply Wiener filter one needs to estimate the covariance matrix (Karhuen-Loeve transform) of the signal. We develop a model for the block covariance structure in the next section.

3. Model of Local Covariance

It is known that wavelet transform has good decorrelating properties and is a reasonable approximation to the Karhuen-Loeve basis of images. Nevertheless, wavelet coefficients are highly structured. Figure 1 shows 512 by 512 fingerprint image and its 256 by 256 horizontal detail wavelet subband.

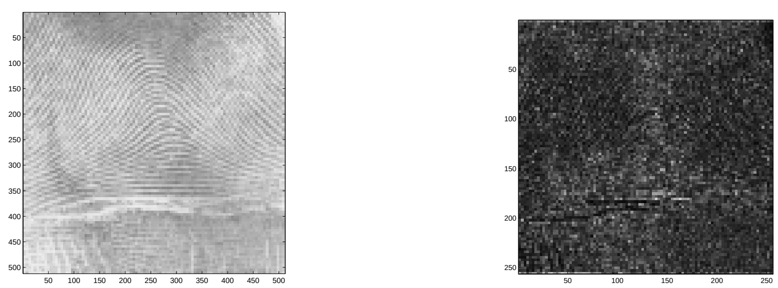


FIGURE 1. Fingerprint image and its horizontal detail wavelet subband.

One can see that large wavelet coefficients (lighter in color) correspond to edges of the original image. Also, because it is the horizontal detail subband, large coefficients tend to align horizontally . It is reasonable to assume that there is

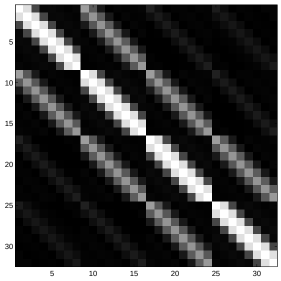


FIGURE 2. Covariance matrix of 4 by 8 block of wavelet coefficients estimated from the horizontal detail subband of the fingerprint image.

correlation within the rows. On the other hand, edges are mostly local and only a few adjacent coefficients should be correlated.

Figure 2 presents covariance matrix of 4 by 8 block of wavelet coefficients estimated from the horizontal detail subband of the fingerprint image. Lighter color corresponds to the elements with larger absolute value. One can see that there is a strong correlation between adjacent coefficients within rows. This confirms our hypothesis.

Figure 2 gives an idea how the block covariance structure of the horizontal detail subband should look like. Nevertheless, it would be an oversimplification to use the same covariance matrix for all blocks. Coefficients are correlated along and decorrelated across the edges. A block without edges should be weakly correlated while a block with many edges should exhibit strong correlation structure of the form similar to the one shown in figure 2. We would like to distinguish between these two types of blocks.

Suppose a vector $\mathbf{X}_j$ corresponds to a block of noisy wavelet coefficients belonging to a given subband. We suggest the following model for its covariance matrix $\mathbf{C}_j$:

$$\mathbf{C}_j = \sigma^2 \mathbf{I} + \gamma_j \mathbf{B}\,. \tag{2}$$

Here σ^2 is the variance of the noise, γ_j is a parameter which has to be estimated from the data, $\mathbf{I}$ is the identity matrix. The covariance matrix $\mathbf{C}_j$ has two parts. $\sigma^2\mathbf{I}$ represents the contribution of the noise (the transform is orthogonal and the wavelet coefficients of the noise are decorrelated). $\gamma_j\mathbf{B}$ is the contribution of the edges. Matrix $\mathbf{B}$ characterizes the local covariance structure of noise-free wavelet coefficients of the particular subband. Figure 2 gives an example of such matrix. We assume that each block $\mathbf{X}_j$ has the same "amount" of noise but different "amount" of edges proportional to γ_j. This parameter changes from block to block.

Parameter γ_j can be estimated using maximum likelihood method. Assuming that $\mathbf{X}_j$ is Gaussian with zero mean and covariance $\mathbf{C}_j$,

$$\mathbf{X}_j \sim \frac{1}{(2\pi)^{n/2}|\mathbf{C}_j|^{1/2}} e^{-\mathbf{X}^T\mathbf{C}_j^{-1}\mathbf{X}/2} = \frac{1}{(2\pi)^{n/2}|\mathbf{C}_j|^{1/2}} e^{-\mathbf{X}^T(\sigma^2\mathbf{I}+\gamma_j\mathbf{B})^{-1}\mathbf{X}/2}$$

we obtain the likelihood functional

$$\begin{aligned} L(\mathbf{Y}_j) = & -\frac{1}{2}\ln|\mathbf{C}_j| - \frac{1}{2}\mathbf{Y}_j(\sigma^2\mathbf{I}+\gamma_j\Lambda)^{-1}\mathbf{Y}_j = \\ & -\frac{1}{2}\sum_{k=1}^{n}\ln(\sigma^2+\gamma_j\lambda_k) - \frac{1}{2}\sum_{k=1}^{n}\frac{y_k^2}{\sigma^2+\gamma_j\lambda_k}\,. \end{aligned} \tag{3}$$

Here $|\mathbf{C}| = \det\mathbf{C}$, y_k, $k = 1,\dots,n$ are components of the vector $\mathbf{Y}_j = \mathbf{Q}\mathbf{X}_j$, λ_k are eigenvalues of $\mathbf{B}$, and $\mathbf{Q}$ is the matrix of eigenvectors of $\mathbf{B}$. Differentiating (3) by γ_j and setting the derivative equal to zero we obtain an equation for γ_j:

$$\frac{1}{\sigma^2}\sum_{k=1}^{n}\frac{\lambda_k y_k^2}{(1+\gamma_j\lambda_k)^2} - \sum_{k=1}^{n}\frac{\lambda_k}{1+\gamma_j\lambda_k} = 0\,. \tag{4}$$

This equation can be solved numerically.

3.1. Estimate of matrix B

In order to solve (4) eigenvalues and eigenvectors of the matrix $\mathbf{B}$ are needed. By our model $\mathbf{B}$ characterizes average covariance structure of the given subband. We estimate it from the mean of $\mathbf{X}_j\mathbf{X}_j^T$ using model (2).

$$\frac{1}{n}\sum_j \mathbf{X}_j\mathbf{X}_j^T = \sigma^2\mathbf{I} + \beta\mathbf{B}\,.$$

Multiplication of $\mathbf{B}$ by a constant changes only normalization and is not important. We set $\mathbf{B}$ to be the difference between $\frac{1}{n}\sum_j \mathbf{X}_j\mathbf{X}_j^T$ and $\sigma^2\mathbf{I}$:

$$\mathbf{B} = \frac{1}{n}\sum_j \mathbf{X}_j\mathbf{X}_j^T - \sigma^2\mathbf{I}\,. \tag{5}$$

4. Denoising Algorithm

Using results from previous sections we suggest the following procedure for image denoising.

1. *Perform orthogonal wavelet decomposition of an image corrupted by Gaussian white noise.*
2. *For each detail (high pass) subband*
   a) *Estimate general block covariance matrix* $\mathbf{B}$ *using (5).*
   b) *Split the subband into non-intersecting blocks* $\mathbf{X}_j$. *Estimate covariance matrix* $\mathbf{C}_j$ *of each block in the form (2). Compute coefficients* γ_j *by solving equations (4) numerically.*
   c) *Apply Wiener filter (1) to each block* $\mathbf{X}_j$ *using covariance matrix* $\mathbf{C}_j$.

3. *Keep scaling (low pass) coefficients unchanged.*
4. *Reconstruct the image from denoised wavelet coefficients.*

We call this algorithm Wiener-wavelet denoising.

5. Numerical Results

We implemented Wiener-wavelet denoising and compared it to two other methods, MATLAB's `wiener2.m` routine and *wavelet soft thresholding.*

`wiener2.m` uses a pixel-wise adaptive Wiener method based on statistics estimated from a local neighborhood of each pixel. In all experiments we chose neighborhood such that the root mean square error (rmse) was minimal.

Value of the threshold in wavelet thresholding was optimized in order to produce the smallest possible rmse.

We used Daubechies least asymmetric orthogonal wavelets with 8 coefficients both in Wiener-wavelet and thresholding methods.

Table 1 compares peak signal to noise ratios (PSNR) obtained by these three methods for different images (PSNR= $20\log_{10}\frac{255}{\text{rmse}}$). Wiener-wavelet denoising always gave the best results.

	Lenna	Barbara	Boats	Yogi
noise	20.16db	20.16db	14.14db	14.14db
thresholding	27.93db	24.00db	23.90db	21.38db
`wiener2.m`	28.13db	24.28db	23.94db	22.18db
Wiener-wavelet	29.53db	25.19db	24.80db	22.81db

TABLE 1. PSNR obtained by different denoising methods.

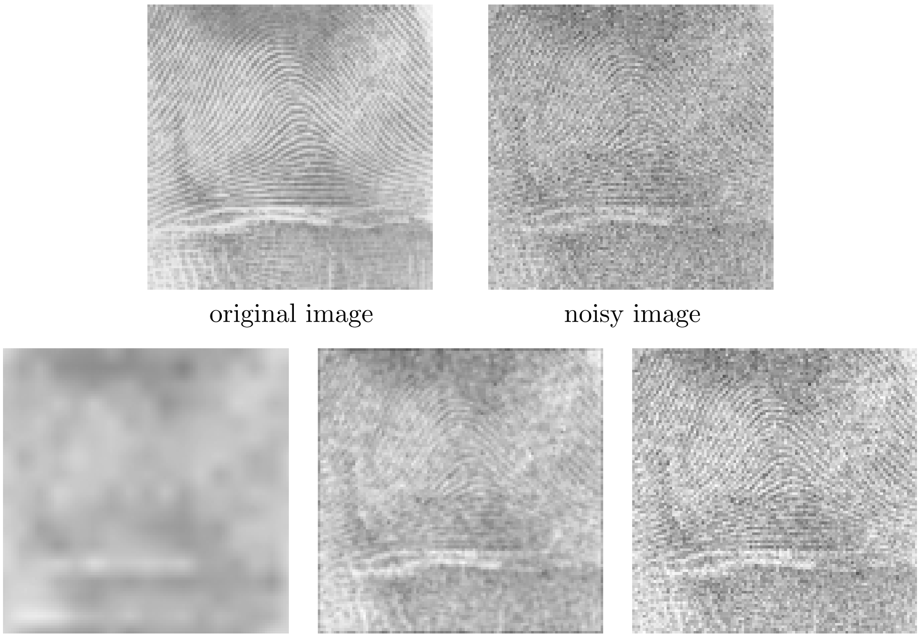
From figure 3 one can see that in the presence of strong noise Wiener-wavelet method preserves edges better. Careful examination shows that some visual information is lost in the noisy image and can be recovered only from Wiener-wavelet denoised image.

6. Conclusions

This paper presents a new technique for image denoising. We assume that blocks of wavelet coefficients are Gaussian with slowly changing covariance matrix. We develop a model for the covariance and use it for Wiener filtering.

In almost all experiments our method produces rmse lower than other methods. Visually Wiener-wavelet method preserves the edges better.

In case of white Gaussian noise with unknown variance Wiener-wavelet denoising allows easy and accurate estimate of the noise variance. If the covariance structure of the noise is know it can be easily taken into account.



original image | noisy image

soft thresholding | `wiener2.m` | Wiener-wavelet method

Figure 3. Visual comparison of different denoising methods.

There is another possible modification of our covariance model. Instead of estimating γ_j as a deterministic parameter one can model it as a random variable. Similar approach is adopted in [8]. We also would like to mention an independent work [5]. There image wavelet coefficients are modeled as a scalar Gaussian variable with spatially changing variance.

Acknowledgements

Author would like to thank Geoffrey Davis and Eero Simoncelli for useful discussions and comments.

References

[1] G. Chang, B Yu and M. Vetterli, *Spatially adaptive wavelet thresholding with context modeling for image denoising*, preprint (1998).

[2] D. Donoho, *De-noising by soft-thresholding*, IEEE Trans. on Info. Theory, **43** (1995), 613–627.

[3] S. Mallat, *Wavelet Tour of Signal Processing*, Academic Press (1998).

[4] S. Mallat, *A theory for multiresolution signal decomposition: the wavelet representation*, IEEE Trans. Patttern Anal. Machine Intell., **11** (1989), 674–693.

[5] K. Mihcak, I. Kozintsev and K. Ramchadndran, *Spatially adaptive statistical modeling of wavelet image coefficients and its application to denoising*, preprint, (1999).

[6] P. Moulin and J. Liu, *Analysis of multiresolution image denoising schemes using a generalized Gaussian and complexity priors*, IEEE Trans. Info. Theory, **45** (1999), 909–919.

[7] E. Simoncelli and E. Adelson, *Noise removal via Bayesian wavelet coring*, Proc. IEEE ICIP, **I** (1996), 379–382.

[8] M. Wainwright and E. Simoncelli, *Scale mixtures of Gaussians and the statistics of natural images* in: S. A. Solla, T. K. Leen, and K.-R. Müller, Eds., Advances in Neural Information Processing Systems 12, (MIT Press, Cambridge MA) (2000).

Department of Mathematics & Computer Science,
Drexel University,
Philadelphia, PA 19104, USA
E-mail address: `vstrela@mcs.drexel.edu`