Published in: Proceedings of the 7th International Conference on Image Processing,
Vancouver, BC, Canada. 10-13 September 2000.

IMAGE DENOISING VIA ADJUSTMENT OF WAVELET COEFFICIENT MAGNITUDE CORRELATION

Javier Portilla and Eero P. Simoncelli

Center for Neural Science, and
Courant Institute of Mathematical Sciences
New York University, NY 10003
{javier,eero}@cns.nyu.edu

We describe a novel method of removing additive white noise of known variance from photographic images. The method is based on a characterization of statistical properties of natural images represented in a complex wavelet decomposition. Specifically, we decompose the noisy image into wavelet subbands, estimate the autocorrelation of both the noise-free raw coefficients and their magnitudes within each subband, impose these statistics by projecting onto the space of images having the desired autocorrelations, and reconstruct an image from the modified wavelet coefficients. This process is applied repeatedly, and can be accelerated to produce optimal results in only a few iterations. Denoising results compare favorably to three reference methods, both perceptually and in terms of mean squared error.

The set of natural images fill a very small fraction of the space of all possible images. Modeling the properties of this set has an enormous importance for many image processing tasks, such as compression or denoising. Typically, such models are statistical, and make use of simplifying assumptions such as stationarity, and spatial localization (e.g., Markov random fields).

In this paper, we consider an image to be a sample of a random field that is parameterized by a small set of statistics. In a very high-dimensional space such as that of all digitized images, the samples of such a random field lie close to the hypersurface of images sharing the same *sample* statistics, and we can approximate the probability density as a uniform distribution over this hypersurface [13]. Thus, assuming the random field parameters are known, the statistical description is replaced by a deterministic one. In this context, an image corrupted by noise is an N-dimensional vector (N the number of pixels) that has been displaced from its original position to a point outside of its associated hypersurface. The problem of estimating the original (noise-free) image involves first estimating the parameters of the hypersurface, and then finding an image on the hypersurface that is close to the observed (noisy) image. Specifically, if one assumes the corrupting noise is Gaussian and white, then the maximum a posteriori (MAP) estimate corresponds to choosing the image on the hypersurface that is closest (in a Euclidean sense) to the observed image. Our method realizes an approximation to this estimate.

1. IMAGE REPRESENTATION

Our model is based on a set of measurements on the coefficients of a multi-scale multi-orientation image representation known as a steerable pyramid [10]. This representation performs a local spectral decomposition of the image using oriented bandpass, self-similar kernels, roughly one octave in bandwidth. It exhibits a number of desirable mathematical properties (it is a tight frame, with translation- and rotation-invariant subbands), and has been used successfully in a number of image processing problems, including noise removal [9]. The use of this representation is also motivated by our knowledge of mammalian visual systems, in which cortical neurons perform a decomposition of the visual input using localized oriented receptive fields. Since human vision is the ultimate criterion of the quality of our processed images, it is desirable to use a set of visually relevant measurements. Recently, we have developed extensions of the steerable pyramid to utilize complex basis kernels, in which the real and imaginary parts are in quadrature phase [7]. Quadrature-pair subbands can be used to detect local features of the image, such as lines and edges, in a spatially shift-invariant way, and they have been widely used both for modeling complex cells in the visual cortex, and in local energy/phase models by the computer vision community. For the results of this paper, we

JP is supported by a fellowship from the Programa Nacional de FPI (Spanish Government). EPS is supported by NSF CAREER grant MIP-9796040.

have used 4 scales and 4 orientation, along with high-pass and low-pass residual subbands.

2. MAGNITUDE CORRELATION

A number of recent denoising methods using joint statistics of the subband coefficients have demonstrated better performance than those based on the wavelet coefficient marginal statistics [e.g., 9, 2, 6, 11]. In particular, we have shown that the magnitudes of pairs of wavelet coefficients are highly correlated [8, 1]. This correlation, mainly caused by features such as lines, edges, and corners, arises between spatial neighbors, and also between coefficients corresponding to different scales and orientations. In a recent work, we used such measurements to capture and reproduce texture [7]. In this paper, we focus only on the spatial correlation (autocorrelation) of the coefficient magnitudes within each subband.

Figure 1(a) shows the magnitude of the output of a complex subband tuned to high frequencies in the vertical direction when applied to a natural image. Note that the large-magnitude coefficients in the subband are primarily arranged in lines following the orientation of the subband. Figure 1(b) shows the central part (in our experiments we have used a 17×17 neighborhood) of the estimated autocorrelation (AC) function. The strong vertical correlation of vertically aligned coefficients is apparent. This behavior is not trivially due to the subband filter, but comes from the special statistics of the natural images: Figures 1(c) and (d) show that adding white noise to the image yields a weaker alignment of the magnitudes.

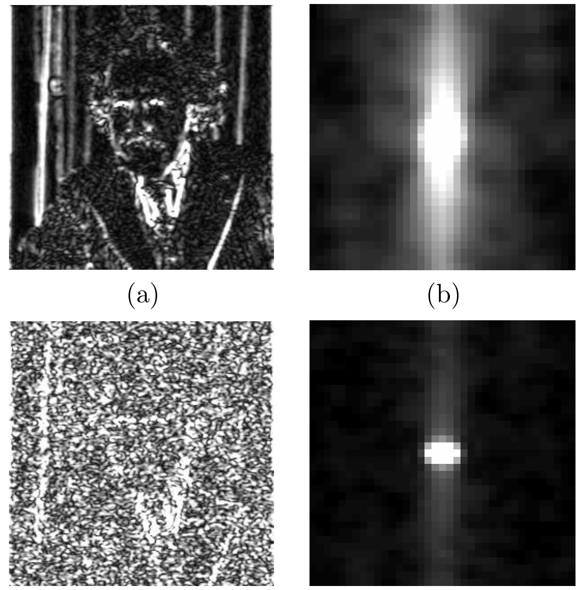
Reversing this reasoning, one would expect to remove a substantial amount of noise in a image corrupted by Gaussian noise by forcing the AC of the coefficient magnitudes in each subband to match the one we would obtain with a noise-free sample (assuming we can estimate such an AC). For the results shown in this paper, we have estimated and adjusted the AC for both the complex subbands and their magnitudes.

3. PARAMETER ESTIMATION

Given a noisy image, and assuming known noise variance, we must estimate the autocorrelation parameters of each subband. The estimation of the original AC of the complex subbands from the observed one is trivially solved by subtracting the AC of the added noise from each subband[1]. But the complex magnitude operation is nonlinear, and an analytic form for the ML estimate of its autocorrelation appears to be intractable. Instead, we have tried to predict the AC functions of the uncorrupted subbands by incorporating a prior. In particular, we have used: 1) the observed normalized AC in the i-th subband, $\mathbf{a_{O,i}}$; 2) the AC corresponding to an input image of univariate Gaussian white noise, $\mathbf{a_{N,i}}$ and 3) a generic (prior) AC, $\mathbf{a_{P,i}}$, obtained by averaging the AC computed from 53 (256×256) normalized reference images. We have observed that the contribution of these last two measurements is roughly proportional to the relative amount of noise, suggesting the following estimator:

$$\mathbf{a}_{E,i} = \mathbf{a}_{O,i} + r_i \left(\lambda_P \mathbf{a}_{P,i} - \lambda_N \mathbf{a}_{N,i} \right),$$

where $r_i = \sigma^2_{N_i} / \sigma^2_{O_i}$, and λ_N and λ_P are chosen to maximize the SNR over the training set. We use a similar form of estimate to compute an overall scale factor for $\mathbf{a}_{E,i}$. This heuristic estimator provides a good fit to the uncorrupted AC training data (SNR of roughly 25 dB). Comparison between denoising results obtained using the AC of the original (clean) image in place of the estimated ones results in only a small improvement in the quality of the cleaned images (roughly 0.3 dB SNR, see Figure 3).



(a) (b) (c) (d)

Figure 1. Wavelet coefficient magnitudes (left) and their autocovariance (right), without (top) and with (bottom) added Gaussian noise. Coefficients are taken from a single vertical subband of a decomposition of the "Einstein" image.

[1] Note that this is not the identity due to the overcompleteness of the Steerable Pyramid.

noisy image	Estimator			
	Linear	*LocWiener*	*Thresh*	*MagCorr*
17.33	25.02	25.48	25.86	26.96
20.53	26.34	27.27	27.45	28.46
24.76	28.31	29.62	29.78	30.29
28.80	30.70	32.05	31.63	32.23

Table 1. Denoising results for four estimators, and four different levels of additive Gaussian noise added to the "Einstein" image. All values indicate peak signal-to-noise ratio (PSNR) in decibels ($20\log_{10}(255/\sigma_{error})$).

4. PROJECTION ALGORITHM

Our method is based on a projection of the observed noisy image onto the surface consisting of all images satisfying the same deterministic constraints as we have estimated for the original. Alternated projections onto constraint sets have been used previously for restoration [12, 5], and also for texture synthesis [4, 7]. For our purposes here, we must impose on every subband and its associated magnitude response the AC parameters we have estimated. Our technique consists of solving for the $M \times N$ kernel $h(\alpha, \beta)$, that satisfies

$$\mathbf{a}_{E,i}(n,m) = \sum_{\alpha,\beta} h(\alpha,\beta)\mathbf{a}_{O,i}(n-\alpha, m-\beta),$$

for the pairs (n, m) within a 17×17 local neighborhood. We then filter the subband in the Fourier domain by multiplying with the absolute value of the square root of the kernel's Fourier transform.

We impose this adjustment in parallel on both the complex signals and their magnitudes for each subband, and then form the estimated coefficients by combining the modified magnitude responses with the phase of the modified complex samples. Finally, we invert the real part of the wavelet pyramid, reconstructing an estimated image. This procedure is then repeated. In order to accelerate the method, on each iteration we amplify the correction made to the estimated image by adding a fraction (specifically 4/5) of the difference between the current and the previous estimated images. The method typically produces optimal results after only two or three iterations. In our MatLab implementation in a standard 300Mhz workstation, denoising of 512×512 images requires roughly three minutes.

5. RESULTS

We have first applied this algorithm to a natural image (not included in our training set), and we have compared the results of our technique with three others: 1) *Linear*: the coefficients are multiplied by an optimal constant for every subband; 2) *LocWiener*, a local Wiener estimator as implemented in the `wiener2` function of the MatLab software package; 3) *Thresh*, optimal threshold application for every subband ([3], as shown in [9]). Figure 2 shows a cropped version of the resulting images. We believe that our model (panel (f)) produces sharper edges and retains more detail, providing a closer resemblance to the original than the other three methods. Table 1 shows the noise level of the processed images (in decibels) as a function of the PSNR level of the input image, for the 4 compared methods.

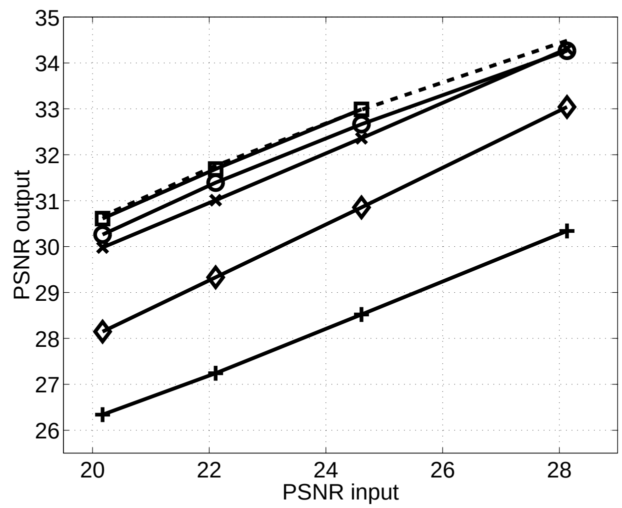


Figure 3. Graphical results comparison (image "Lenna"). From top to bottom. Dashed line: Upper bound for *MagCorr* (no estimation error). Squares: *SAWT* using an over-complete representation [2]. Circles: *MagCorr*. Xs: *5x5 LAWMAP* [6]. Diamonds: *locWiener* (MatLab's `wiener2` function). Pluses: *Thresh* [3].

Finally, we have compared the performance of our method with two state-of-the-art techniques [2, 6] using the image "Lenna", with the same set of noise variance values used in [6]. Figure 3 shows a plot of these results, where the abscissa represents the PSNR of the noisy image and the ordinate the PSNR of the denoised image. We have also included in this graphic (dashed line) the upper bound performance of our method assuming known original AC constraints, as well as the results obtained with the *Thresh* and *LocWiener* methods explained before. Our own results (upper continuous line with circles) lie in between those of [2] and [6]. Next in performance is *LocWiener* and the last one is *Thresh*. Considering the simplicity of the proposed method, we believe these results are very encouraging for exploring other forms of denoising via constraint projection.

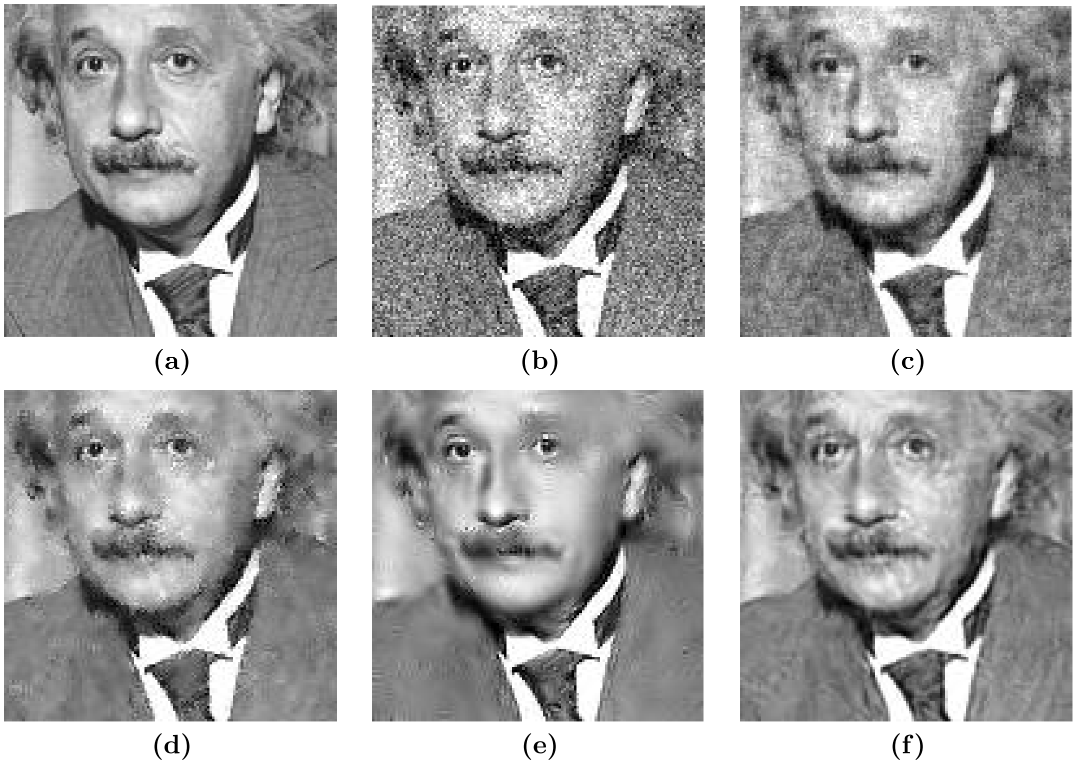


Figure 2. **(a)**: Original "Einstein" image (cropped). **(b)**: Noisy image (PSNR 20.5dB). **(c)**: Linear least-squares estimator (26.3dB). **(d)**: Local Wiener estimation (27.3dB). **(e)**: Optimal thresholding (27.4dB). **(f)**: Magnitude correlation model (28.46dB).

6. REFERENCES

[1] R. W. Buccigrossi and E. P. Simoncelli. Image compression via joint statistical characterization in the wavelet domain. *IEEE Trans Image Proc*, 8(12):1688–1701, December 1999.

[2] S. G. Chang, B. Yu, and M. Vetterli. Spatially adaptive wavelet thresholding with context modeling for image denoising. In *Fifth IEEE Int'l Conf on Image Proc*, Chicago, October 1998.

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

[4] D. Heeger and J. Bergen. Pyramid-based texture analysis/synthesis. In *Proc. ACM SIGGRAPH*, August 1995.

[5] A. N. Hirani and T. Totsuka. Dual domain interactive image restoration: Basic algorithm. In *Third Int'l Conf on Image Proc*, volume I, pages 797–800, 1996.

[6] M. K. Mihçak, I. Kozintsev, K. Ramchandran, and P. Moulin. Low–complexity image denoising based on statistical modeling of wavelet coefficients. *IEEE Trans. on Signal Processing*, 6(12):300–303, December 1999.

[7] J. Portilla and E. Simoncelli. A Parametric Texture Model based on Joint Statistics of Complex Wavelet Coefficients Int'l Journal of Computer Vision. 2000. To appear.

[8] E. P. Simoncelli. Statistical models for images: Compression, restoration and synthesis. In *31st Asilomar Conf on Signals, Systems and Computers*, pages 673–678, Pacific Grove, CA, November 1997.

[9] E. P. Simoncelli. Bayesian denoising of visual images in the wavelet domain. In P. Müller and B. Vidakovic, editors, *Bayesian Inference in Wavelet Based Models*, chapter 18, pages 291–308. Springer-Verlag, New York, June 1999. Lecture Notes in Statistics, vol. 141.

[10] E. P. Simoncelli and W. T. Freeman. The steerable pyramid: A flexible architecture for multi-scale derivative computation. In *Second Int'l Conf on Image Proc*, volume III, pages 444–447, Washington, DC, October 1995.

[11] V. Strela. Denoising via block wiener filtering in wavelet domain. In *3rd European Congress of Mathematics*, Barcelona, July 2000. Birkhäuser Verlag.

[12] D. C. Youla. Generalized image restoration by the method of alternating orthogonal projections. *IEEE Trans. Circuits and Systems*, 25:694–702, 1978.

[13] S. C. Zhu, X. Liu, and Y. N. Wu. Exploring the Julesz ensemble by efficient Markov chain monte carlo. In *IEEE Workshop on Statistical and Computational Theories of Vision*, Fort Collins, CO, 1999.