

WAVELET DENOISING BY RECURSIVE CYCLE SPINNING

Alyson K. Fletcher and Kannan Ramchandran

University of California, Berkeley
Department of Electrical Engineering
Berkeley, CA 94720 USA
{alyson,kannanr}@eecs.berkeley.edu

Vivek K Goyal

Digital Fountain
39141 Civic Center Drive, Suite 300
Fremont, CA 94538 USA
v.goyal@ieee.org

ABSTRACT

Coupling the periodic time-invariance of the wavelet transform with the view of thresholding as a projection yields a simple, recursive, wavelet-based technique for denoising signals. Estimating a signal from a noise-corrupted observation is a fundamental problem of signal processing which has been addressed via many techniques. Previously, Coifman and Donoho introduced *cycle spinning*, a technique estimating the true signal as the linear average of individual estimates derived from wavelet-thresholded translated versions of the noisy signal. Here, it is demonstrated that such an average can be dramatically improved upon. The proposed algorithm recursively “cycle spins” by repeatedly translating and denoising the input via basic wavelet denoising and then translating back; at each iteration, the output of the previous iteration is used as input. Exploiting the convergence properties of projections, the proposed algorithm can be regarded as a sequence of denoising projections that converge to the projection of the original noisy signal to a small subspace containing the true signal. It is proven that the algorithm is guaranteed to globally converge, and simulations on piecewise polynomial signals show marked improvement over both basic wavelet thresholding and standard cycle spinning.

1. INTRODUCTION

The basic principle of wavelet denoising is to identify and zero out wavelet coefficients of the signal which are likely to contain mostly noise. By identifying and preserving significant coefficients, wavelet thresholding preserves important highpass features of the signal such as discontinuities. This property is useful, for example, in image denoising to maintain the sharpness of edges in the image.

Recently, Coifman and Donoho [5] suggested an improvement to basic wavelet thresholding known as *cycle spinning*. It is known that the wavelet transform is periodically time-invariant. By using different shifts of the noisy signal, one can compute different estimates of the unknown signal, and then linearly average the estimates. The idea is that the estimates at different shifts have statistically different errors, which will be reduced by averaging.

In this work, the cycle spinning algorithm is re-examined in the context of projections. Wavelet thresholding, in essence, identifies a subspace in which the unknown signal is likely to have most of its energy and projects the noisy signal onto this subspace. Cycle spinning thereby identifies the subspaces, performs the corresponding denoising projections, and averages the projections.

However, instead of averaging the projections, it can be argued that it is better to project once onto the *intersection* of the sub-

spaces. The reason is simple. If the unknown signal has most of its energy in each of the subspaces, it will have most of its energy in their intersection. Since the intersection of subspaces typically has a much smaller dimension than the subspaces themselves, projecting to the intersection produces an estimate much closer to the true signal.

To perform the projection onto the intersection of the subspaces, a novel algorithm called *recursive cycle spinning* is proposed. The algorithm recursively produces a sequence of estimates by repeatedly cycling through the shifts, using the result of one iteration as the input for the subsequent iteration. It is proved analytically that the estimates from the proposed method globally converge.

A simple numerical simulation of the algorithm for denoising a piecewise polynomial signal is presented. In the simulation, recursive cycle spinning outperforms both basic wavelet thresholding and standard cycle spinning.

2. WAVELET DENOISING TECHNIQUES

2.1. Basic Denoising

Consider estimating an unknown discrete-time signal $x[n]$, $n \in [0, N - 1]$ from a noise-corrupted signal,

$$y[n] = x[n] + d[n], \quad (1)$$

where $d[n]$ is additive noise. Let $X[n] = \mathcal{W}(x[n])$ denote the discrete wavelet transform of $x[n]$. Only orthogonal wavelet transforms are considered here.

Basic wavelet denoising is performed by taking the wavelet transform of the noise-corrupted $y[n]$ and then zeroing out the detail coefficients that fall below a certain threshold. An inverse wavelet transform is applied to the thresholded signal to yield an estimate for the true signal, as below:

$$\hat{x}[n] = D(y[n]) = \mathcal{W}^{-1}(\Lambda_t(\mathcal{W}(y[n])))$$

where Λ_t is the diagonal thresholding operator that zeroes out wavelet coefficients less than the threshold, t .

Denoising by wavelet thresholding was introduced by Donoho and Johnstone in [7]. Their method of using “soft thresholding” was referred to as “wavelet shrinkage.” The idea of setting fine-scale wavelet coefficients to zero as a noise reduction method was not revolutionary. However, earlier and contemporaneous works on this method, *e.g.* [12], emphasized the need to separate “edge” and “non-edge” locations and threshold only the fine-scale coefficients in non-edge locations. Techniques for this separation were

effective but decidedly ad hoc. Donoho and Johnstone's wavelet shrinkage was remarkable for its simplicity, requiring no tracking or correlation of maxima and minima across scales.

2.2. Cycle Spinning

One way of improving upon basic wavelet denoising is through cycle spinning, where the signal to be denoised is translated by various time shifts. Cycle spinning utilizes the periodic time-invariance of the wavelet transform. The wavelet transform is not time-invariant. Consequently, if the noisy signal is shifted in time, denoised, and then shifted back, the result will, in general, be different from the estimate obtained from denoising without shifting. Assuming a periodic invariance with period M , for $i \in [0, M-1]$, $\mathcal{W}(x[n])$ and $\mathcal{W}(x[n+i])$ are related in a complicated manner.

For a sequence $y[n]$, let $y^{(i)}[n] = y[n+i]$ denote a right-shifted version. In cycle spinning, for each $i \in [0, M-1]$, the noisy signal $y^{(i)}[n]$ is denoised via wavelet thresholding. Each of these denoised outputs is left-shifted by i , and these M estimates are denoted $\hat{x}^{(i)}[n]$ with $i \in [0, M-1]$. The cycle spinning estimate is obtained by simply linearly averaging the M estimates, $\hat{x}^{(i)}[n]$, $i = 0, \dots, M-1$.

The errors in the estimates are not completely dependent. Consequently averaging these estimates yields a reduction in noise.

3. RECURSIVE CYCLE SPINNING

It is natural to wonder if linear averaging is the best way to utilize the information from these M estimates. Averaging is typically far from optimal. A simple example is presented illustrating that significant improvement should be possible.

3.1. Motivating Example

Consider estimating a piecewise constant signal $x[n]$ from a noise corrupted signal $y[n]$. It is natural to use a wavelet transform with one vanishing moment; in this case, the 1-stage Haar filter bank with filters

$$H_0(z) = \frac{1}{\sqrt{2}}(1 + z^{-1}), \quad H_1(z) = \frac{1}{\sqrt{2}}(1 - z^{-1})$$

Since this is a one-stage filter bank, cycle spinning would entail averaging estimates obtained from shifts of zero and one. In this example, the detail wavelet coefficients are the output of $H_1(z)$ downsampled by two. Applying this transform to the piecewise constant signal $x[n]$ yields detail coefficients of

$$X_1^{(0)}(k) = \frac{1}{\sqrt{2}}(x[2k] - x[2k-1]).$$

Similarly, the transform leads to detail coefficients for the shifted signal of

$$X_1^{(1)}[k] = \frac{1}{\sqrt{2}}(x[2k+1] - x[2k]).$$

For $x[n]$ constant on some interval, say $[0, M]$, the detail coefficients $X_1^{(0)}(k)$ and $X_1^{(1)}(k)$ are zero for $k \in [0, M/2]$. Now consider $y[n]$, the signal $x[n]$ corrupted by noise. Let $\hat{x}^{(i)}$ denote the estimates from the i^{th} shift, for $i \in \{0, 1\}$. If the noise is sufficiently small, the detail coefficients for $k \in [0, M/2]$ of the noisy signal will be zeroed out via thresholding. Consequently, the

estimates obtained from denoising the shifts of the signal will have the properties that

$$\hat{x}^{(0)}[2k] = \hat{x}^{(0)}[2k-1] \text{ and } \hat{x}^{(1)}[2k+1] = \hat{x}^{(1)}[2k].$$

Therefore on the interval of discussion, the estimate on each shift will be constant over pairs of samples.

Thus with small noise assumed, denoising with no shift reveals that the signal is constant over odd pairs of samples, and denoising with the shift of one reveals constancy over even pairs. Simultaneously utilizing both sets of information yields a resulting estimate for $x[n]$ that is constant over the entire interval $[0, M]$. Simply linearly averaging the two estimates will actually produce an estimate where no consecutive pair of coefficients are in general constant. This motivates seeking an alternative method for combining the estimates.

3.2. Basic Algorithm

Basic wavelet denoising can be fundamentally viewed as a projection. Uniting this view with the periodic time-invariance of the wavelet transform, the M different estimates used in cycle spinning are actually obtained from M different denoising projections. Thus, an improved estimate can be formed by projecting the noisy signal $y[n]$ onto the intersection of the range spaces of these M projections, denoted the *denoised space intersection*. Unfortunately, there is no simple way to compute this projection operator from the M projection operators from the different shifts.

The proposed algorithm, *recursive cycle spinning*, is an approximate iterative algorithm. The algorithm generates a sequence of estimates, \hat{x}_ℓ . The algorithm is initialized to the original noisy signal, $\hat{x}_0[n] = y[n]$, and recursively updated with the rule

$$\hat{x}_{\ell+1} = D_i(\hat{x}_\ell), \quad i = \ell \bmod M, \quad (2)$$

where $D_i(\cdot)$ is the denoising operator that uses a shift of i . The algorithm repeatedly cycles through the M shifts, using the denoised output from one iteration as the input to the next. The basic idea is that the algorithm tries to converge to a fixed point. Any fixed point, \hat{x}_∞ , will satisfy

$$\hat{x}_\infty = D_i(\hat{x}_\infty)$$

for all i , and will therefore be in the range space of denoising projections for all i , as desired.

3.3. Windowed Thresholding

Recursive cycle spinning requires particular attention to the thresholding. Discontinuities in a signal will, in general, result in a number of non-zero detail coefficients depending on the filter length. However, some of these coefficients may be small and could be zeroed out with standard thresholding. This could often even result in an increase in the SNR. However, in recursive cycle spinning, if too many of these coefficients are eliminated, the recursive projections eventually smooth across the discontinuity to make all the wavelet coefficients zeros. To avoid this problem, we employed a *windowed thresholding* method where not only are the wavelet coefficients above the threshold level zeroed out, but also those close to a coefficient that is above the threshold. This window size is dependent both upon the filter length and the stage of the filter bank. Windowed thresholding allows for the final estimate to have a sufficient number of non-zero detail coefficients around a discontinuity.

4. CONVERGENCE ANALYSIS

An obvious issue for the proposed recursive cycle spinning algorithm is convergence. A complicating factor in the analysis is that even for the same shift value i , different coefficients may be zeroed out by different iterations. Therefore, the denoising operations will not necessarily correspond to the same projection for each iteration. Also, one can imagine the different projections “bouncing” between the range spaces without converging. Despite these complications, it is proven that the recursive cycle spinning algorithm does indeed converge.

Theorem 1 *Let \hat{x}_ℓ be a sequence of signals produced by the recursive cycle spinning algorithm (2) with some initial condition $\hat{x}_0 = y$. Suppose that the denoising operations are performed with an orthogonal wavelet transform, and assume that the threshold levels are continuous functions of the input. Then \hat{x}_ℓ converges to a signal \hat{x}_∞ with the property that*

$$\hat{x}_\infty = D_i(\hat{x}_\infty)$$

for all shifts i .

The theorem shows that the recursive cycle spinning algorithm globally converges. The proof is based on a lengthy projection analysis, which is omitted for space considerations. The fundamental idea in the proof is that it can be shown that projecting repeatedly from a finite set of projections, without consideration of order, but with each projection occurring infinitely often, converges to the projection onto the intersection of the range spaces of the projections.

Since the convergence point \hat{x}_∞ satisfies $\hat{x}_\infty = D_i(\hat{x}_\infty)$ for all i , \hat{x}_∞ converges to a *denoised space intersection*. Theorem 1 does not relate the convergence point, \hat{x}_∞ , of the algorithm to the “true” signal. For this analysis, further assumptions are required.

Suppose that the noisy signal $y[n]$ is given by (1) for some signal $x[n]$. Let $X^{(i)}(k)$ denote the k^{th} wavelet coefficient of $x[n]$ left-shifted by i . Assume that for all shifts i , each wavelet coefficient $X^{(i)}(k)$ satisfies either

- a) $X^{(i)}(k) = 0$, or
- b) $|X^{(i)}(k)| > T(k)$.

The assumption can be understood as follows: Suppose the true signal $x[n]$ were used as the input to wavelet thresholding. Since $X^{(i)}(k) = 0$ for any k with $|X^{(i)}(k)| \leq T(k)$, the thresholding would not eliminate any non-zero wavelet coefficients. Consequently, if D_i is the wavelet thresholding operator with a shift of i , the assumption is equivalent to

$$x = D_i(x)$$

for all i .

Now let Z_i be the set of coefficients k such that $X^{(i)}(k) = 0$ and let R_i be the set of signals $\hat{x}[n]$ whose wavelet transform coefficients satisfy

$$\hat{X}^{(i)}(k) = 0, \quad \forall k \in Z_i.$$

By definition, the true signal $x[n] \in R_i$ for all i . For any shift i , wavelet thresholding on the noisy signal $y[n]$ would ideally zero out exactly the wavelet coefficients Z_i and project $y[n]$ onto R_i . As discussed earlier, the goal of the algorithm is to project the noisy signal $y[n]$ onto the intersection of the spaces R_i . The following theorem states that if the noise is sufficiently small, recursive cycle spinning does indeed converge to the correct projection.

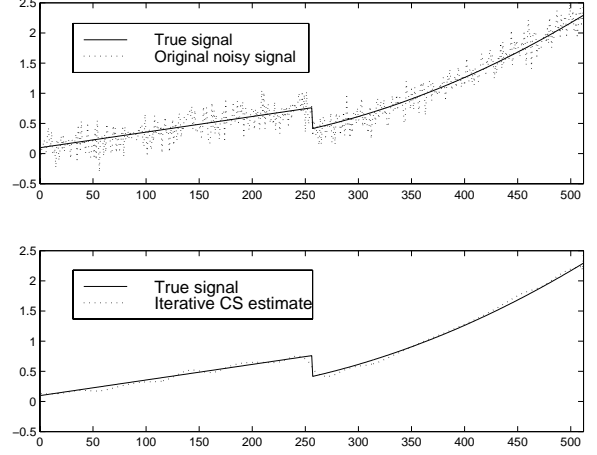


Fig. 1. A typical result of recursive cycle spinning estimation. Top panel: True piecewise polynomial signal and the original noise-corrupted signal from which the estimate is made (SNR = 15 dB). Bottom panel: Final estimate from recursive cycle spinning (SNR = 27.9 dB).

Theorem 2 *Suppose, in addition to the assumptions of Theorem 1, the noisy signal y is given by (1) for a signal x satisfying the assumptions above. Then, there exists an $\epsilon > 0$ such that if $\|d\| < \epsilon$, \hat{x}_ℓ converges to the projection of y onto the intersection of the spaces R_i .*

Theorem 2 does not state that the recursive cycle spinning estimates converge to the true signal x itself. This is, in general, impossible with finite noisy data. Instead, the theorem shows that recursive cycle spinning converges to an intersection of a number of spaces containing the true signal, and therefore the final estimate is close to the true signal.

5. NUMERICAL RESULTS

This section provides a few examples to demonstrate the convergence of the recursive cycle spinning algorithm and the improved performance over denoising in a single basis and over conventional cycle spinning.

Consider the problem of estimating

$$x[n] = \begin{cases} 0.08n + 3 & 0 \leq n < 256 \\ 0.004n^2 - 0.08n + 7 & 256 \leq n < 512 \end{cases} \quad (3)$$

from the noise-corrupted signal

$$y[n] = x[n] + d[n],$$

where $d[n]$ is zero-mean, white, Gaussian noise. The variance of $d[n]$ is set so that the original SNR is 15 dB. Since $x[n]$ is a piecewise polynomial signal with degree at most two, wavelet thresholding techniques will be effective as long as the highpass filters of the wavelet transform have at least three vanishing moments. Thus, for the initial experiment the recursive cycle spinning algorithm was applied with a two-stage wavelet transform implemented with D_3 filters.

The wavelet thresholding was implemented with windowed thresholding described in Section 3.3. The thresholding levels, T , were computed separately for each subband using the simple heuristic rule $T = 3\sigma$, where σ is the root mean square value of the wavelet coefficients in that subband. This level ensures a high probability that any wavelet coefficient with only noise is zeroed out.

The top panel of Fig. 1 shows the desired signal $x[n]$ and the noisy signal $y[n]$. The bottom panel shows the result of 400 iterations of recursive cycle spinning, which improves the SNR to 27.9 dB and gives a visually close estimate. As the algorithm progresses, the MSE decreases monotonically. It drops fastest at the beginning and the estimate is within 1 dB of the final error after 100 iterations. It takes only a few iterations for the MSE to reach the level of -22.7 dB obtained with standard cycle spinning.

For a more extensive comparison between basic wavelet thresholding, cycle spinning, and recursive cycle spinning, a length-1024 signal similar to (3) was denoised many times with several wavelet transforms. The Daubechies D_3 and D_4 filter pairs were used in wavelet transforms with 1, 2, 3, and 4 stages. In these experiments, the number of iterations for the recursive cycle spinning algorithm is 10 times the number of shifts.

The results of 50 trials with each wavelet transform are shown in Fig. 2. Recursive cycle spinning consistently outperforms standard cycle spinning by 3–4 dB, while cycle spinning outperforms basic wavelet thresholding by about 1 dB. As expected, the noise reduction is greater with the D_3 filter pair because its highpass filter has the minimum required number of vanishing moments. Experiments with input SNRs ranging from -6 dB to 22 dB gave very similar results.

6. REFERENCES

- [1] A. Chambolle, R. A. DeVore, N.-y. Lee, and B. J. Lucier. Nonlinear wavelet image processing: Variational problems, compression, and noise removal through wavelet shrinkage. *IEEE Trans. Image Proc.*, 7(3):319–335, March 1998.
- [2] S. G. Chang, B. Yu, and M. Vetterli. Adaptive wavelet thresholding for image denoising and compression. *IEEE Trans. Image Proc.*, 9(9):1532–1546, September 2000.
- [3] S. G. Chang, B. Yu, and M. Vetterli. Spatially adaptive wavelet thresholding with context modeling for image denoising. *IEEE Trans. Image Proc.*, 9(9):1522–1531, September 2000.
- [4] S. G. Chang, B. Yu, and M. Vetterli. Wavelet thresholding for multiple noisy image copies. *IEEE Trans. Image Proc.*, 9(9):1631–1635, September 2000.
- [5] R. R. Coifman and D. L. Donoho. Translation-invariant de-noising. In A. Antoniadis and G. Oppenheim, editors, *Wavelets and Statistics*, volume 103 of *Springer Lecture Notes in Statistics*, pages 125–150, New York, 1995. Springer-Verlag.
- [6] D. L. Donoho. De-noising by soft-thresholding. *IEEE Trans. Inform. Th.*, 41(3):613–627, May 1995.
- [7] D. L. Donoho and I. M. Johnstone. Ideal spatial adaptation via wavelet shrinkage. *Biometrika*, 81:425–455, 1994.
- [8] P. L. Dragotti and M. Vetterli. Wavelet transform footprints: Catching singularities for compression and denoising. In *Proc. IEEE Int. Conf. Image Proc.*, volume 2, pages 363–366, Vancouver, September 2000.
- [9] H. Krim, D. Tucker, S. Mallat, and D. Donoho. On denoising and best signal representation. *IEEE Trans. Inform. Th.*, 45(7):2225–2238, November 1999.
- [10] X. Li and M. T. Orchard. Spatially adaptive image denoising under overcomplete expansion. In *Proc. IEEE Int. Conf. Image Proc.*, volume 3, pages 300–303, Vancouver, September 2000.
- [11] M. K. Mihçak, I. Kozintsev, K. Ramchandran, and P. Moulin. Low-complexity image denoising based on statistical modeling of wavelet coefficients. *IEEE Sig. Proc. Lett.*, 6(12):300–303, December 1999.
- [12] Y. Xu, J. B. Weaver, D. M. Healy, Jr., and J. Lu. Wavelet transform domain filters: A spatially selective noise filtration technique. *IEEE Trans. Image Proc.*, 3(6):747–758, November 1994.

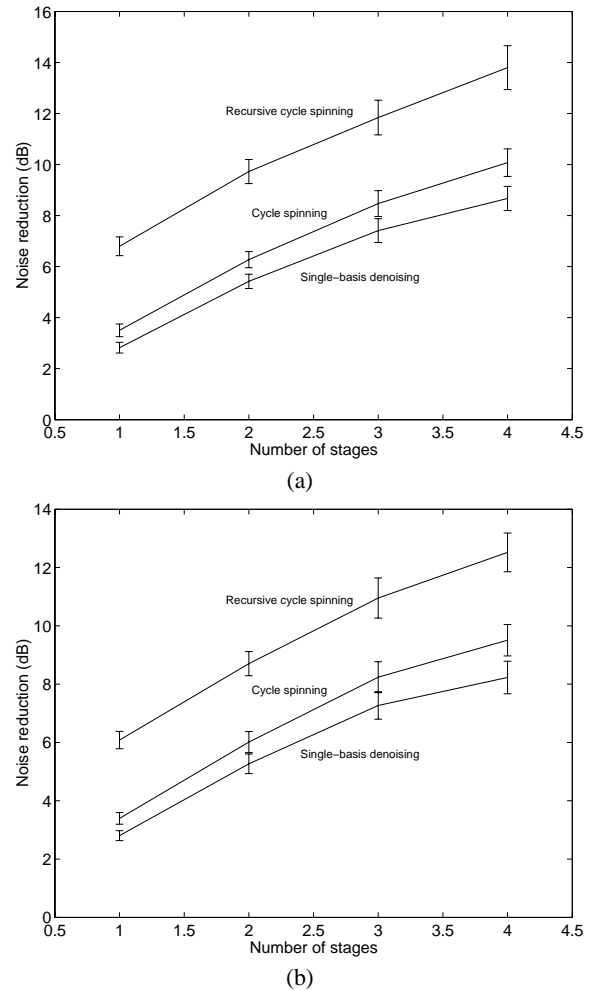


Fig. 2. Comparison of (single-basis) wavelet thresholding, cycle spinning, and recursive cycle spinning for denoising a piecewise quadratic signal. For each technique and number of wavelet transform stages, the means and standard deviations of 50 experiments are shown. (a) D_3 filter pair; (b) D_4 filter pair.