ON MULTISCALE WAVELET ANALYSIS FOR STEP ESTIMATION

Brian M. Sadler and Ananthram Swami

Army Research Laboratory, 2800 Powder Mill Road, Adelphi, MD 20783-1197

ABSTRACT

We consider step detection and estimation using a multiscale wavelet analysis, based on the ability of a certain discrete wavelet transform (DWT) to characterize signal steps and edges. This DWT, developed by Mallat and Zhong, estimates the gradient at various smoothing levels without downsampling in time. As first proposed by Rosenfeld for edge sharpening, multiple scales are combined by forming the pointwise product across scales. We show that this approach is a non-linear whitening transformation, and characterize the non-Gaussian pdf of the output. Detection curves are shown for parameterized sigmoidal step change signals. Step location estimation performance is also shown, with comparison to the Cramer-Rao bound in additive white Gaussian noise.

1. INTRODUCTION

Edge and step detection and estimation is a fundamental problem in many areas of signal and image processing. A simple general approach, that works well at high SNR, is gradient estimation. Many of the basic image edge detection operators (e.g., Roberts, Prewitt, Sobel) reduce in one-dimension to an FIR filter with response [-1, 0, 1]. More general extensions, so-called filtered derivative methods, combine smoothing with gradient estimation to reduce noise effects. This results in classical detection and estimation tradeoffs, e.g., between the level of smoothing and the variance of the estimated step location. The problems with the level of smoothing appropriate for gradient estimation can be overcome to some extent by employing a multi-scale analysis, i.e., combining results over multiple levels of smoothing.

It is well known that wavelets may be used for singularity detection [1]. This has been applied to edge detection by Mallat and Zhong [2], putting earlier work of Canny and others into the WT framework. Consider a continuous-time wavelet $\psi(t)$ that consists of the first derivative of a smoothing function $u(t)$, given by $\psi(t) = du(t)/dt$. For a signal of interest $x(t)$ it is straightforward to show that [1]

$$W_s x(t) = x(t) * \left(s \frac{du_s}{dt} \right)(t) = s \frac{d}{dt}(x * u_s)(t), \qquad (1)$$

where $u_s(t) = (1/s)u(t/s)$ and $*$ denotes convolution, provided that $\psi(t)$ is a valid wavelet with zero first moment. Thus, for appropriate choice of $u(t)$, $W_s x(t)$ can be interpreted as a derivative of a local average of $x(t)$ where the degree of smoothing depends on scale s. The result is derivative estimation at various levels of smoothing.

We consider a scheme based on the ability of a certain discrete wavelet transform (DWT) to characterize the local regularity of signals. In [2] Mallat and Zhong developed a DWT based on a discrete-time approximation to $u(t)$ using a cubic spline; implementation details are in [2]. Note that this DWT is dyadic in scale but not in time; there is no downsampling of the filter outputs. The IR's for the first few scales are shown in Figure 1, and the resulting frequency responses are shown in Figure 2. Note the linear regions in Figure 2 correspond to differentiation at the various smoothing levels.

Hereafter we restrict our attention to Mallat and Zhong's DWT, although our results are general for a family of linear derivative estimation filters. Let the DWT input be $x(n)$, with jth scale output denoted

$$y_j(n) = \sum_k h_j(k) x(n-k), \qquad (2)$$

where $h_j(n)$ is the IR of the jth DWT filter. In the following we consider a multiscale detection and estimation strategy based on the product of the DWT outputs, given by

$$p(n) = \prod_{j=j_0}^{j_1} W_{2^j} x(n) = \prod_{j=j_0}^{j_1} y_j(n). \qquad (3)$$

This approach was first suggested (before the advent of wavelets) by Rosenfeld [5], who employed dyadic smoothing scales and where the smoothing filter taps were all equal to one. This approach has recently been employed by Xu et al for image denoising [7].

An example of (3) is given in Figure 3, depicting a time series, its DWT for the first three scales, and normalized $p(n)$ for $j_0 = 1$ and $j_1 = 2$ and 3. The signal is two-valued in white Gaussian noise. The peaks in the DWT at various scales correspond to the edges. Because the peaks align across the first few scales, the product $p(n)$ exhibits corresponding peaks. Peaks do not align across arbitrarily high scales because neighboring peaks interfere due to lengthening filter responses. The peaks in $p(n)$ are generally well pronounced, except for those corresponding to the isolated impulses between $n = 400$ and 450 in the original time series, where smoothing leads to weakened response at the higher scales. Peaks in Figure 3e are generally positive going because of the even number of products, whereas those in Figure 3f are bipolar and preserve the edge up/down direction information. In the following we develop properties and analyze behavior of $p(n)$.

2. CROSS-SCALE CORRELATION

In this section we consider the cross-correlation between outputs of the DWT. We assume that $x(n)$ is real-valued iid with correlation function $E[x(n)x(n+m)] = \sigma_x^2\delta(m)$. Then

$$E[y_i(n)y_j(n)] = \sigma_x^2 \sum_k h_i(k)h_j(k). \tag{4}$$

Thus, the correlation coefficient between DWT outputs is given by

$$r_{ij} = \frac{\sum_k h_i(k)h_j(k)}{(\sum_{k_1} h_i^2(k_1) \sum_{k_2} h_j^2(k_2))^{1/2}}, \tag{5}$$

and the joint covariance matrix of the DWT outputs at time n has (i,j) element given by

$$[C]_{ij} = \sigma_x^2 \sum_k h_i(k)h_j(k). \tag{6}$$

These relations are easily calculated for the specific values of the FIR DWT filters $h_i(k)$. Correlation coefficients r_{ij} are shown in Table 1 for $i,j = 1,\cdots,5$.

1.0000	0.5345	0.2097	0.0759	0.0270
	1.0000	0.6444	0.2791	0.1074
		1.0000	0.6787	0.3019
			1.0000	0.6868
				1.0000

Table 1: Correlation coefficients for DWT outputs $y_i(n)$.

3. A NONLINEAR WHITENING TRANSFORMATION

Let $p_K(n) = y_1(n)\cdots y_K(n)$ be the Kth product of the outputs of the DWT, corresponding to $j_0 = 1$ and $j_1 = K$ in (3). For simplicity in the following we restrict ourselves to $j_0 = 1$, but the extension to arbitrary j_0 is straightforward. For $x(n)$ zero-mean Gaussian, we find the autocorrelation of $p_K(n)$ as follows. Consider

$$\begin{aligned} r_{p_K}(\tau) &= E[p_K(n)p_K(n+\tau)] \qquad (7)\\ &= E[\prod_{i=1}^{K}(\sum_{l_i} h_i(l_i)x(n-l_i)) \prod_{j=1}^{K}(\sum_{m_j} h_j(m_j)x(n+\tau-m_j))]\\ &= \sum_{l_1} h_1(l_1)\cdots\sum_{l_K} h_K(l_K) \sum_{m_1} h_1(m_1)\cdots\sum_{m_K} h_K(m_K)\\ &\times E[x(n-l_1)\cdots x(n-l_K)x(n+\tau-m_1)\cdots x(n+\tau-m_K)]. \end{aligned}$$

The expectation is a $(2K)$th joint moment that reduces to a sum of $1\times 3\times\cdots\times(2K-1)$ terms, where each term is composed of appropriate products of $r_{ij}(0)$ and $r_{ij}(\tau)$, defined as

$$r_{ij}(\tau) = \sum_l h_i(l)h_j(l+\tau). \tag{8}$$

For example, for $K = 2$ the expectation in (7) reduces to 3 terms, and using the whiteness of $x(n)$ we have

$$r_{p_2}(\tau) = \sigma_x^4[r_{12}^2(0) + r_{11}(\tau)r_{22}(\tau) + r_{21}(\tau)r_{12}(\tau)]. \tag{9}$$

Similarly, for $K = 3$, the expectation reduces to a sum of 15 product terms in $r_{ij}(\tau)$.

We note that the case of $x(n)$ non-Gaussian can be handled similarly by expressing the moment in (7) in terms of cumulants and exploiting the assumed iid nature of $x(t)$. This will also involve higher-order analogs of (8).

For Mallat's DWT $r_{p_K}(m)$ may be calculated for various cases, although the number of terms grows quickly for $K > 3$. Plots of $r_{p2}(\tau)/r_{p2}(0)$ and $r_{p3}(\tau)/r_{p3}(0)$ are shown in Figure 4. $r_{p3}(\tau)$ exhibits a strong zero-lag peak with small values elsewhere. This effect becomes more pronounced as K is increased, as confirmed by simulations. Despite both the temporal dependence in each of the DWT time series $y_k(n)$, as well as the cross-correlation between these outputs, $p_K(n)$ is effectively whitened for $K \geq 3$. This is intuitively apparent from Figure 2, because the time domain product corresponds to convolution in the frequency domain.

4. DETECTION

Although it is whitened with respect to its second-order statistics, $p(n)$ is in general distinctly non-Gaussian. However, determination of the first-order pdf is not straightforward even with $x(n)$ Gaussian, except when $K = 2$ (e.g., see [3, sect. 2.3]). For third or higher order products we must resort to numerical techniques, such as nonparametric kernel-based pdf estimation. Figure 5 shows estimated pdf's with $x(n)$ white Gaussian for two cases, $K = 2$ (top) and $K = 3$ (bottom). Cauchy and unit-variance normal pdf's are shown for reference. Note the skewness in the $K = 2$ case, due to the positive correlation between $y_1(n)$ and $y_2(n)$, and because K is even. In this case the positive tail is heavier than Cauchy, showing a strong impulsive nature in $p_2(n)$ on the positive side, with a much lighter tail on the negative side. In the $K = 3$ case the pdf is symmetric with tail behavior heavier than Gaussian but lighter than Cauchy. An alternative point of view is to regard $p_K(n)$ as an instantaneous estimate of a cross-correlation. For example, $p_3(n)$ corresponds to a triple-correlation; we immediately conclude that if $x(n)$ is zero-mean Gaussian then in this case $E[p_3(n)] = 0$.

Figure 6 shows detection results, where detection is declared by comparing $|p_3(n)|$ with a threshold for each n. The threshold was obtained from the estimated pdf for $p_3(n)$, and set to achieve $P_{fa} = 0.001$. Three signals were used; two were based on a sigmoidal step model given by

$$s(n) = \frac{m_2 + m_1\exp\{-\alpha(nT - n_0T - \tau)\}}{1+\exp\{-\alpha(nT - n_0T - \tau)\}}, \tag{10}$$

where m_1 and m_2 are the signal levels before and after the step, T is the sampling interval time, parameter α determines the rise time, and the step occurs (in continuous time) at $n_0T + \tau$. Here τ is uniformly distributed in $[0,T]$, modeling the effects of time quantization via sampling. The SNR is defined locally as

$$\text{SNR} = \frac{(m_2 - m_1)^2}{\sigma^2}, \tag{11}$$

where σ^2 is the noise variance, Two values were used, $\alpha = 4.4$, and $\alpha = 1$, corresonding to a fast and slower step rise-

time. In both cases $m_1 = 0$ and $T = 1$, and m_2 was adjusted to achieve the desired SNR. The third signal in Figure 6 was an ideal step function (a rise-time of 1 sample), from $m_1 = 0$ to m_2, where m_2 is set to achieve the desired SNR. As we would expect, slowing rise-time results in poorer detection performance.

5. ESTIMATION

Next we consider estimation of step change location. This assumes that a change has occured in the observation interval, hence there is no detection problem. Reza and Doroodchi [4] have derived CRB's for the case of sigmoidal step change as in (10) in additive white Gaussian noise, where the SNR is defined as in (11). The CRB depends primarily on data local to the change time, in a neighborhood approximately given by $|n - n_0| \leq 2.5t_r$, where n_0 is the change time and t_r is the step change rise time (time to change from 10% to 90% of the step height).

Figure 7 depicts the theoretical CRB, as well as experimental MSE, for two methods. The two estimation methods are first, based on $p_3(n)$, and second, based on a simple gradient estimator with FIR given by [-1, 0, 1]. The sigmoidal function was generated with $m_1 = 0$, $T = 1$, and $\alpha = 4.4$, corresonding to a rise-time of about $4T$ (see [4]). The step height m_2 was set to achieve the desired SNR. The simple gradient estimator has essentially no smoothing, whereas $p_3(n)$ exploits multiple smoothing levels. At low SNR the benefit of the smoothing is evident as the DWT method outperforms the simple gradient estimator by about 5 dB. At high SNR both methods are time-quantization error dominated. The simple gradient approaches the minimum variance of 1/12 of a sampling time, which arises due to the uniform distribution of the sampling phase error. The DWT method approaches 1/2, reflecting the effects of smoothing via the increased variance of the estimate.

Changing α changes the effective rise-time of the sigmoidal signal. Decreasing α creates a slower rising step. This results in a shift upward of the CRB as the step location is inherently more difficult to estimate. Repeating the experiment in Figure 6 with lower α results in a shift to the right as higher SNR is needed to obtain the same performance; while the general behavior of the curves is the same.

6. CONCLUSIONS

We have analyzed a method for nonlinearly combining multiscale wavelet outputs for detecting and estimating steps and edges. The smoothed gradient DWT developed by Mallat provides a wavelet framework for this approach, as this DWT is dyadic in scale but not in time. The cross-scale product approach was originally proposed by Rosenfeld for edge sharpening. Despite its non-linear nature the cross-scale product is whitening, but the resulting noise pdf is distinctly non-Gaussian and generally heavy tailed. The pdf was characterized and detection results presented parameterized by step rise-time. Location estimation was also characterized and compared to the Cramer-Rao bound in white Gaussian noise. The results indicate tradeoffs possible in exploiting multiple smoothing levels when the most appropriate single smoothing level is not known a priori.

7. REFERENCES

[1] S. Mallat and W. L. Hwang, "Singularity detection and processing with wavelets," IEEE Trans. Info. Theory, Vol. 38, pp. 617–643, 1992.

[2] S. Mallat and S. Zhong, "Characterization of signals from multiscale edges," IEEE Trans. Pattern Anal. and Mach. Intell., Vol. 14, No. 7, pp. 710–732, July 1992.

[3] K. S. Miller, Multidimensional Gaussian Distributions (Wiley, 1964).

[4] A. M. Reza and M. Doroodchi, "Cramer-Rao lower bound on locations of sudden changes in a steplike signal," IEEE Trans. SP, Vol. 44, No. 10, pp. 2551–2556, 1996.

[5] A. Rosenfeld, "A nonlinear edge detection technique," IEEE Proc., pp. 814–816, May 1970.

[6] B. M. Sadler, T. Pham, and L. C. Sadler, "Optimal and wavelet-based shockwave detection and estimation," J. Acoust. Soc. Amer., to appear; see also Proc. Intl. Conf. Acoust., Speech, and Sig. Proc. (ICASSP97), Vol. 3, pp. 1889–1992, Munich, Germany, April 1997.

[7] Y. Xu, J. Weaver, D. Healy, and J. Lu, "Wavelet transform domain filters: a spatially selective noise filtration technique," IEEE Trans. on Image. Proc., Vol. 3, No. 6, pp. 747–758, November 1994.

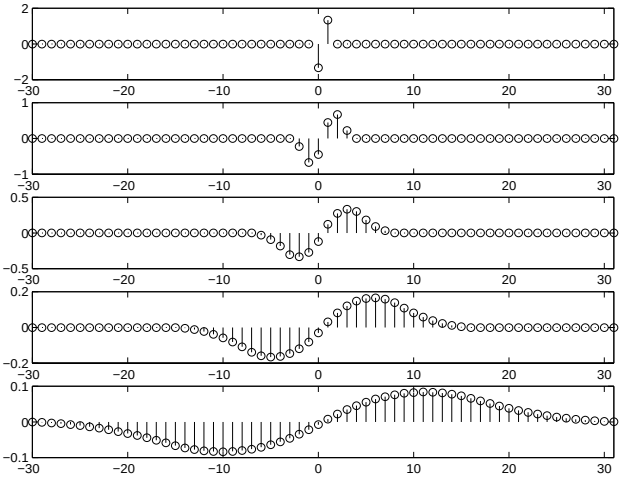

Figure 1. Impulse responses of the DWT for the first five scales.

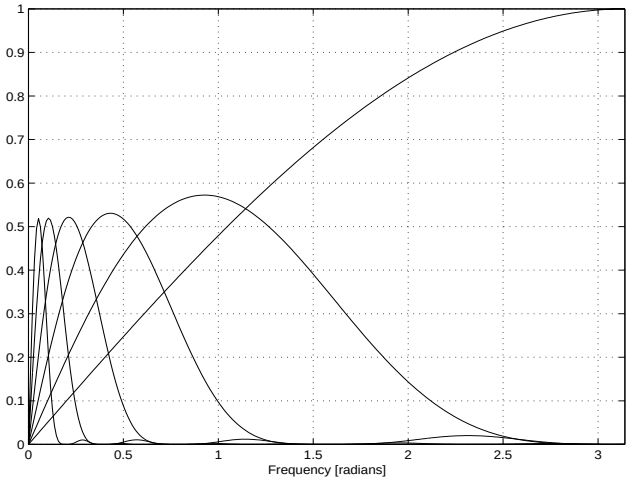

Figure 2. Frequency responses of the DWT for six scales.

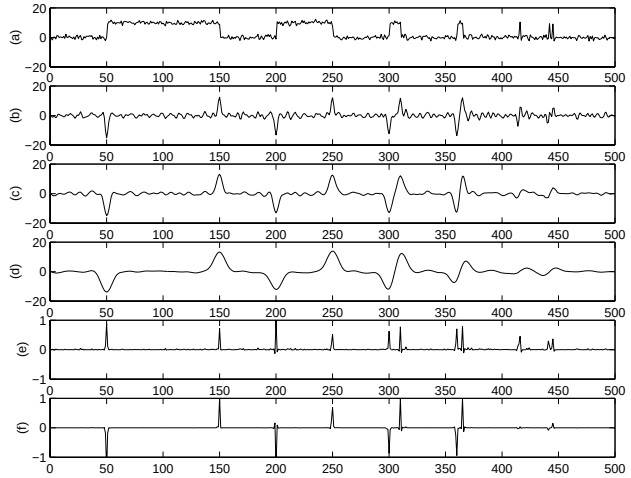

Figure 3. DWT example showing (a) time series, (b)-(d) first 3 scales of the DWT, (e) $p_2(n)$, (f) $p_3(n)$.

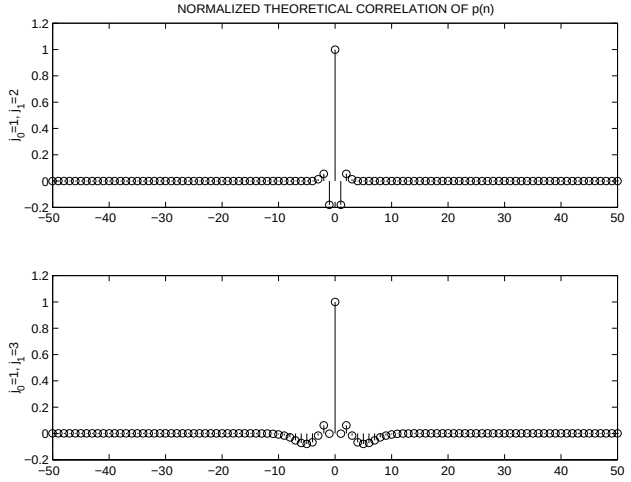

Figure 4. Theoretical normalized correlations of $p_2(n)$ and $p_3(n)$.

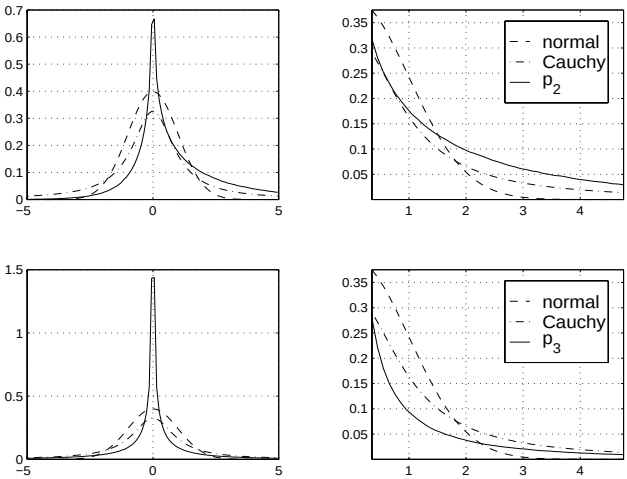

Figure 5. Estimated first-order pdfs of $p_2(n)$ and $p_3(n)$, with normal and Cauchy also shown.

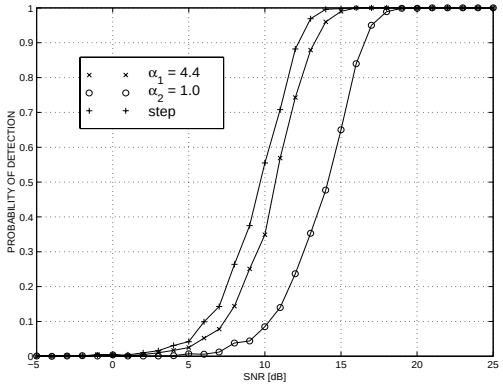

Figure 6. Probability of step detection using $p_3(n)$ in white Gaussian noise.

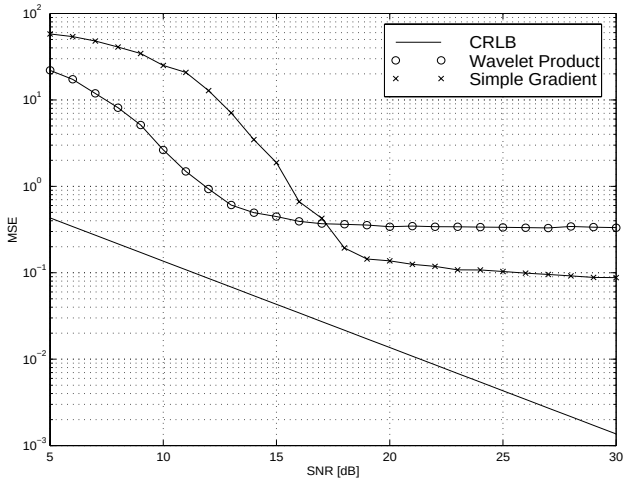

Figure 7. MSE's and theoretical CRB for step change location estimation.