AN UNDECIMATED WAVELET TRANSFORM BASED DETECTOR FOR TRANSIENTS IN 1/f NOISE

Thomas T. Liu and Antony C. Fraser-Smith

Space, Telecommunications and Radioscience Laboratory
Stanford University, Stanford, California 94305, U.S.A.
e-mail: ttliu@alpha.stanford.edu

ABSTRACT

We consider the detection in the presence of $1/f$ noise of a known transient signal of unknown amplitude, scale and delay. We introduce a generalized likelihood ratio test (GLRT) method based on pattern matching in the undecimated discrete wavelet transform (UDWT) domain. In many cases, the computational complexity of the detector can be reduced with minimal performance impact by limiting the pattern matching operations to locations in the UDWT domain that correspond to the existence of transform local maxima. As examples of our approach, we simulate the detection of transients that are modeled either by scaling functions, Gaussian functions, or two-sided exponential functions.

1. INTRODUCTION

Wavelet transforms have been widely applied to the problem of transient detection and processing, primarily because the transform basis functions provide good time localization [1, 2, 3]. A number of detection methods have been proposed that involve the tracking of local transform maxima across analysis scales [3, 4]. These techniques rely on the observation that the evolution of the transform maxima across scales provides a measure of the local regularity of the signal [3]. Other proposed detectors have been based on standard detection theory. Frisch and Messer [2] observed that the wavelet transform acts as bank of matched filters, and can therefore be used as a generalized likelihood ratio test (GLRT) detector for transients modeled by wavelets in the presence of white noise. For transients of unknown shape, GLRT detectors that form the detection statistic in the wavelet domain have also been discussed [1, 5]. For a transient with known parameters in $1/f$ Gaussian noise, Wornell [6] described a matched filter detector in the Discrete Wavelet Transform (DWT) domain. This detector relies on the observation that the DWT acts as an approximate whitening transform for $1/f$ noise.

In this paper we introduce a GLRT detector for a known transient signal of unknown amplitude, scale and delay parameters in $1/f$ Gaussian noise. Such a detector is widely applicable because $1/f$ power spectra have been observed in a broad range of physical processes [6].

The proposed method utilizes pattern matching in the Undecimated Discrete Wavelet Transform (UDWT) domain and is the shift invariant extension of the scheme presented in [6]. The computational cost of the method can be significantly reduced by limiting the matching operations to those locations in the UDWT where the local maxima propagate across scales. Monte Carlo simulations are used to compute the receiver operating characteristics (ROC) for transients modeled by scaling functions, Gaussian functions, and two-sided exponential functions. The simulation results indicate that the reduction in computational complexity can be achieved with negligible impact on detector performance. Futhermore, this work helps to explain how maxima tracking algorithms can be understood within the framework of detection theory.

Partial support for this work was provided by the U.S. Geological Survey through Grants no. 1434-HQ-97-GR-03124 and 1434-HQ-96-GR-02715. Additional funding for the first author was provided by the generous support of a Gerald J. Lieberman Fellowship.

2. NOTATION

We use the notation $\psi(t)$ to represent wavelets and the subscript notation $\psi_m(t) = 2^{-m/2}\psi(2^{-m}t)$ and $\psi_{m,n}(t) = 2^{-m/2}\psi(2^{-m}t - n)$ for wavelets at other scales. The inner product of two functions is defined as $\langle f(t), g(t)\rangle = \int_{-\infty}^{\infty} f(t)g^*(t)dt$ where g^* is the complex conjugate of g.

The DWT of a signal $x(t)$ is $X_{m,n} = \langle x(t), \psi_{m,n}(t)\rangle$. The choice of time origin for the basis functions $\psi_{m,n}(t)$ is arbitrary, and we define other DWT's with basis functions $\psi_{m,n}(t-J)$ and the notation

$$X^{[J]}_{m,n} = \langle x(t), \psi_{m,n}(t-J)\rangle. \quad (1)$$

If M denotes the largest analysis scale of interest, then the $X^{[J]}_{m,n}$ are invariant to shifts by integer multiples of 2^M, i.e. $X^{[J]}_{m,n} = X^{[J-2^M]}_{m,n+2^{(M-m)}}$. As a result there are 2^M unique DWT shifts, with each shift giving rise to a different decomposition of the signal $x(t)$. The UDWT is $\tilde{X}_{m,n} = \langle x(t), \psi_m(t-n)\rangle$. Note that the UDWT is a shift invariant transform, i.e. $x(t-j) \Rightarrow \tilde{X}_{m,n-j}$.

3. THEORY

We consider the detection of a transient with unknown amplitude, scale, and delay. We have the standard hypothesis test

$$\begin{array}{rcl} H_0: \; x(t) & = & v(t) \\ H_1: \; x(t) & = & As_{k,l}(t) + v(t) \end{array}$$

where A, k, and l are the unknown amplitude, scale, and delay, respectively, of the transient, and we assume that $A > 0$ and k and l are integers. We define $s_{k,l}(t) = 2^{\gamma k/2}\xi_k(t-l)$ where $\xi_k(t) = 2^{-k/2}\xi(2^{-k}t)$ and $\xi(t)$ is the signal model. The additive noise $v(t)$ is assumed to be a $1/f^\gamma$ Gaussian random process.

A standard scheme for detection with unknown parameters is the generalized likelihood ratio test GLRT [2, 7]. It has the form: choose H_1 if the likelihood ratio $r(x(t)) > r_1$ where

$$r(x(t)) = \frac{\max_{\{A,k,l\}} p\,(x(t)|A,k,l,H_1)}{p\,(x(t)|H_0)}$$

and r_1 is a threshold value chosen to achieve a desired probability of false alarm (PFA), and choose H_0 otherwise.

Detection in $1/f$ noise is based on the observation by Wornell and others ([6] and references therein) that the DWT acts as an approximate whitening transform for $1/f^\gamma$ noise processes. If $v(t)$ represents the noise process with power spectrum $\frac{\sigma_v^2}{f^\gamma}$, then the DWT coefficients $V_{m,n}$ are approximately uncorrelated with variance $\sigma_w^2 2^{\gamma m}$ where $\sigma_w^2 = \eta\sigma_v^2$ and η is a function of the wavelet and parameter γ.

We consider first the detection problem with known signal parameters in order to develop some ideas that will be useful in understanding the unknown parameter case. Following Wornell [6], we note that the equivalent hypothesis test in the DWT domain is

$$\begin{array}{rcl} H_0: \; X^{[0]}_{m,n} & = & V^{[0]}_{m,n} \\ H_1: \; X^{[0]}_{m,n} & = & AS^{[0],\{k,l\}}_{m,n} + V^{[0]}_{m,n} \end{array} \qquad (2)$$

where A, k and l are known parameters, $X^{[J]}_{m,n}$ was defined in (1), and $S^{[J],\{k,l\}}_{m,n} = \langle s_{k,l}(t), \psi_{m,n}(t-J)\rangle$. The likelihood ratio is

$$\frac{p\left(\mathbf{X}^{[0]}|H_1\right)}{p\left(\mathbf{X}^{[0]}|H_0\right)} = \frac{\prod_{m,n}\exp\left(-\frac{\left(X^{[0]}_{m,n} - AS^{[0],\{k,l\}}_{m,n}\right)^2}{2\sigma_w^2 2^{\gamma m}}\right)}{\prod_{m,n}\exp\left(-\frac{\left(X^{[0]}_{m,n}\right)^2}{2\sigma_w^2 2^{\gamma m}}\right)} \qquad (3)$$

where $\mathbf{X}^{[J]} = \{X^{[J]}_{m,n}\}$ is the vector of observations.

We can simplify (3) to obtain the sufficient test statistic $\lambda(\mathbf{X}) = \sum_{m,n} 2^{-\gamma m} X^{[0]}_{m,n} S^{[0],\{k,l\}}_{m,n}$, where we have dropped the subscript on $\mathbf{X}$ for notational convenience. Because the DWT is shift variant, the performance of the detector is also shift variant. To see this, we consider the performance index $d = \frac{A\sqrt{\mathcal{E}_k(l)}}{\sigma_w}$ which is the normalized distance between the distributions of $\lambda(\mathbf{X})$ under the two hypotheses in (2). For each parameter value k, $\mathcal{E}_k(l) = \frac{1}{A}E\left[\lambda(\mathbf{X})|H_1\right] = \sum_{m,n} 2^{-\gamma m}\left(S^{[0],\{k,l\}}_{m,n}\right)^2$ where $E[\lambda(\mathbf{X})|H_1]$ is the expected value of $\lambda(\mathbf{X})$ given hypothesis H_1. Figure 1 shows an example of the variation of $\mathcal{E}_k(l)$ with input shift l when $\xi(t)$ is a Coiflet parameter 2 (henceforth denoted as C_2) scaling function [8] and $\gamma = 1$.

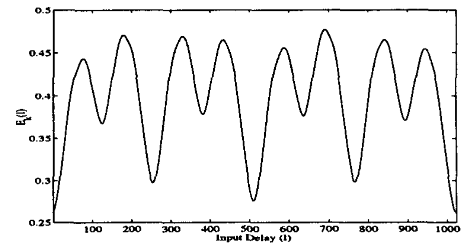

Figure 1: Variation of $\mathcal{E}_k$ with input shift for a C_2 scaling function ($k = 7$) analyzed with a C_2 wavelet over the range $m = 1$ to 10 with $\gamma = 1$.

We define a shift invariant detector by first noting that the noise statistics are independent of DWT shift, since all shifted transforms also act as approximate whitening filters. As a result, for a given known signal $s_{k,l}(t)$, we are free to choose the DWT shift that maximizes $\mathcal{E}_k(l)$ and thus the performance of the detector. For the signal $s_{k,0}(t)$ with delay $l = 0$, we define a detector with a test statistic of the form $\lambda(\mathbf{X}) = \sum_{m,n} 2^{-\gamma m} X^{[J_k]}_{m,n} S^{[J_k],\{k,0\}}_{m,n}$ where $J_k = \arg\max_J \sum_{m,n} 2^{-\gamma m}\left(S^{[J],\{k,0\}}_{m,n}\right)^2$ is the optimum DWT shift and is a function of k. It can be shown [9] that $\sum_{m,n} 2^{-\gamma m}\left(S^{[J_k],\{k,0\}}_{m,n}\right)^2 = \max_l \mathcal{E}_k(l)$. We introduce the notation $\mathcal{E}_k = \max_l \mathcal{E}_k(l)$ for use in the remainder of the paper.

For the case of arbitrary delay l we define a test statistic $\lambda(\mathbf{X}, l) = \sum_{m,n} 2^{-\gamma m} X^{[J_k(l)]}_{m,n} S^{[J_k(l)],\{k,l\}}_{m,n}$ where $J_k(l) = \arg\max_J \sum_{m,n} 2^{-\gamma m}\left(S^{[J],\{k,l\}}_{m,n}\right)^2$. With the identity $S^{[J+l],\{k,l\}}_{m,n} = S^{[J],\{k,0\}}_{m,n}$, we find that

$$\sum_{m,n} 2^{-\gamma m}\left(S^{[J_k],\{k,0\}}\right)^2 = \sum_{m,n} 2^{-\gamma m}\left(S^{[J_k+l],\{k,l\}}\right)^2,$$

so that $J_k(l) = J_k + l$. We may therefore rewrite the test statistic as $\lambda(\mathbf{X}, l) = \sum_{m,n} 2^{-\gamma m} X^{[J_k+l]}_{m,n} S^{[J_k],\{k,0\}}_{m,n}$. The corresponding likelihood ratio is $\frac{p(\mathbf{X}^{[J_k+l]}|H_1)}{p(\mathbf{X}^{[J_k+l]}|H_0)}$.

We use the concepts developed above to define a generalized likelihood ratio for the case of unknown parameters:

$$r(\tilde{\mathbf{X}}) = \max_{\{A,k,l\}} \frac{p\left(\mathbf{X}^{[J_k+l]}|A,k,l,H_1\right)}{p\left(\mathbf{X}^{[J_k+l]}|H_0\right)}.$$

where $\tilde{\mathbf{X}} = \{\tilde{X}_{m,n},\; m,n \in \mathcal{Z}\}$ is the vector of UDWT observations. Note that the DWT shift $J_k + l$ is chosen to maximize the detector performance conditioned on the unknown parameters k and l. The ratio simplifies to yield a sufficient test statistic

$$\lambda(\tilde{\mathbf{X}}) = \max_{\{A,k,l\}}\left[A\sum_{m,n} 2^{-\gamma m} X^{[J_k+l]}_{m,n} S^{[J_k],\{k,0\}}_{m,n} - \frac{A^2}{2}\sum_{m,n} 2^{-\gamma m}\left(S^{[J_k],\{k,0\}}_{m,n}\right)^2\right].$$

From the definition of J_k the second term is equal to $\frac{A^2}{2}\mathcal{E}_k$. We show in [9] that the definition $s_{k,l}(t) = 2^{\gamma k/2}\xi_k(t)$ results in $\mathcal{E}_k = \mathcal{E}_0$ for all k, where $\mathcal{E}_0$ is a constant. Thus we may write

$$\lambda(\tilde{\mathbf{X}}) = \max_{\{A,k,l\}} \left[A \sum_{m,n} 2^{-\gamma m} X_{m,n}^{[J_k+l]} S_{m,n}^{[J_k],\{k,0\}} - \frac{A^2}{2}\mathcal{E}_0 \right].$$

The maximization over k and l can be performed independently of the maximization over A, and we may define a test statistic

$$\begin{aligned} \lambda(\tilde{\mathbf{X}}) &= \max_{\{k,l\}} \lambda(\tilde{\mathbf{X}}, k, l) \\ &= \max_{\{k,l\}} \sum_{m,n} 2^{-\gamma m} X_{m,n}^{[J_k+l]} S_{m,n}^{[J_k],\{k,0\}} \qquad (4) \\ &= \max_{\{k,l\}} \sum_{m,n} 2^{-\gamma m} \tilde{X}_{m,2^m n+J_k+l} \tilde{S}_{m,2^m n+J_k}^{\{k,0\}}. \qquad (5) \end{aligned}$$

4. IMPLEMENTATION

Equations (4) and (5) represent two equivalent ways of computing the detection statistic. We implement (5) which can be viewed as a pattern matching operation in the UDWT domain, where at each scale the pattern is $\tilde{S}_{m,2^m n+J_k}^{\{k,0\}}$. This procedure requires $O\left(2N\sum_{k\in\mathcal{K}} P_k\right)$ operations, where N is the length of the signal, P_k is the number of non-zero coefficients in $\tilde{S}_{m,2^m n+J_k}^{\{k,0\}}$ and $\mathcal{K}$ is the set of unknown scales. The contribution of most of the coefficients to the detection process is, however, negligible. We reduce the number of required coefficients by ranking them according to their contribution to $\mathcal{E}_k$ and selecting only the largest T_k coefficients. In the case of the C_2 scaling function and $\gamma = 1$, we find that the largest $T_k = 20$ coefficients account for 99.4 percent of the value of $\mathcal{E}_k$.

We can further reduce the number of computations by observing that the most significant coefficients of the pattern $\tilde{S}_{m,2^m n+J_k}^{\{k,0\}}$ tend to be located near peaks in the UDWT domain. Figure 2 shows an example for the C_2 scaling function and the Gaussian function. This is not a surprising result since the shift J_k was chosen to maximize $\mathcal{E}_k$. In particular, we note that a number of the coefficients lie somewhere on the maximum peaks at each analysis scale m. We expect, therefore, that the locations in the UDWT domain where the local maxima propagate across scales are likely to be the locations at which the detection statistic is maximized. Adapting the nomenclature of [10], we refer to these locations as ***transform ridges*** and constrain the computation of $\lambda(\mathbf{X}, k, l)$ to a subset of values that correspond to the ridges. Typically, the size of this subset is one to two orders of magnitude less than the size N of the original signal. The preceding discussion leads us to propose the following detection methods:

Method A. *Baseline GLRT method.* Compute $\lambda(\mathbf{X}, k, l)$ for all values of k and l and find the maximum.

Method B. *Preselection of Scales and Delays using Transform Ridges.* In this method we use a ridge finding algorithm to generate estimates of both the scale and location of the transient. The qth estimate is referred to as an ordered pair $\{\hat{k}_q, \hat{l}_q\}$. We compute the detection statistic $\lambda(\mathbf{X}) = \max_q \lambda(\mathbf{X}, \hat{k}_q, \hat{l}_q)$.

Method C. *Preselection of Delays using Transform Ridges.* In this method we use a ridge finding algorithm to estimate the delays but not the scales. The detection statistic is $\lambda(\mathbf{X}) = \max_{k,q} \lambda(\mathbf{X}, k, \hat{l}_q)$.

The ridge finding algorithm is described in detail in [9], where it is also shown that the computational complexity of methods B and C are typically an order of magnitude less than that of method A. There are many variations of the proposed methods, some of which we describe in [9].

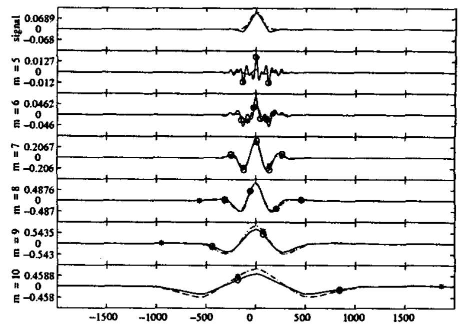

Figure 2: The UDWT of a C_2 scaling function (solid line) and the UDWT of a Gaussian function (dashed-dotted line) are shown for scales $m = 5$ to 10. The signals are shown in the top subfigure ($k = 7$ for both cases). The largest 20 coefficients (as ranked by contribution to $\mathcal{E}_k$, with $\gamma = 1$) of the pattern $\tilde{S}_{m,2^m n+J_k}^{\{k,0\}}$ are marked with 'o' and '*' symbols for the scaling function and Gaussian function, respectively.

5. SIMULATION RESULTS

We used Monte Carlo simulations to obtain the receiver operating characteristics (ROC) of the detectors described above. Each simulation consisted of 200 independent trials. We generated $1/f^\gamma$ noise sample paths using the method described in [11] and $\gamma = 1$. We employed three different transient signal models: a C_2 scaling function, a Gaussian function $g(t) = \alpha e^{-at^2}$, and a two-sided exponential function $e(t) = \beta e^{-b|t|}$. The temporal width and parameter $\mathcal{E}_k$ were made identical for all signal models. We also defined the functions $g_s(t) = (\sqrt{2})^{(\gamma-1)} g(\sqrt{2}t)$ and $e_s(t) = (\sqrt{2})^{(\gamma-1)} e(\sqrt{2}t)$ for use in examining the effects of scale mismatch. For each signal type, we used $P_k = 20$. The set of unknown input scales was $\mathcal{K} = \{6, 7, 8\}$, and the length of each signal was $N = 16384$. We computed the UDWT using a C_2 wavelet for analysis scales $m = 1$ to 10 and used the performance index $d = \frac{A\sqrt{\mathcal{E}_k}}{\sigma_w}$ as a measure of the signal to noise ratio of each simulation. Note that for a GLRT detector, d no longer represents the normalized distance between the distributions under the two detection hypotheses.

In Figure 3 we show ROC curves for a C_2 transient at input scales 6 through 8. The performance of method C is equivalent to that of method A, while that of method B

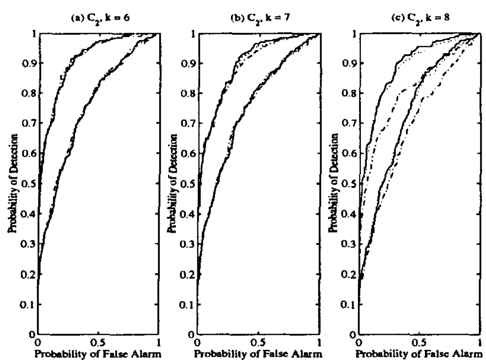

Figure 3: ROC curves for detector methods A (solid line), B (dashed-dotted line), and C (dashed line). Input signal is C_2 scaling function at scales ranging from 6 to 8. The upper and lower set of curves correspond to $d = 5$ and $d = 4$, respectively.

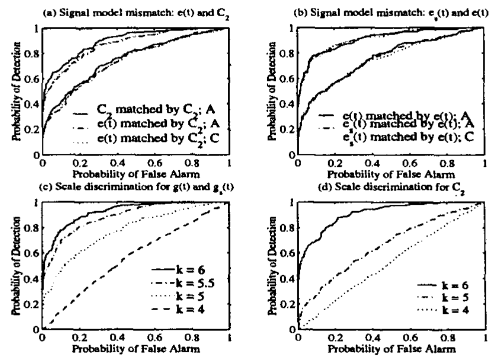

Figure 4: ROC curves: signal model mismatch in panel (a), scale mismatch in panel (b), and scale discrimination in panels (c) and (d). For panels (a) and (b) the input scale of the transients is $k = 7$ and the upper and lower set of curves correspond to $d = 5$ and $d = 4$, respectively. For panels (c) and (d), $d = 5$.

is worse for some cases, indicating errors in the estimation of the unknown scale. In Figure 4 panel (a) we present the ROC curves for a C_2 transient detected with method A using a C_2 UDWT signal pattern and for a $e(t)$ transient detected with methods A and C, also with a C_2 pattern. The plot shows that the detector is relatively insensitive to small errors in the assumed signal model. In Figure 4 panel (b) we show the ROC curves for a $e(t)$ transient detected with method A using a $e(t)$ UDWT signal pattern and for a $e_s(t)$ transient detected with methods A and C, also with a $e(t)$ pattern. A $e_s(t)$ transient with scale $k = 7$ has an effective scale of 6.5, and therefore panel (b) shows that the detector is fairly robust with respect to mismatch between the actual input scale and the assumed input scales. Finally, in Figure 4 panels (c) and (d) we examine the ability of the detector to discriminate against transients with scales outside the assumed scale range $\mathcal{K} = \{6, 7, 8\}$. We consider Gaussian transients with scales $4, 5, 5.5,$ and 6 and C_2 transients with scales $4, 5,$ and 6. For both cases we find that by scale $k = 4$ the performance is that of a detector which assumes no *a priori* information about the signal.

6. CONCLUSION

We have proposed a GLRT detector for transients in $1/f^\gamma$ noise by making use of the approximate whitening properties of the DWT. The detector is shift invariant and is implemented with a pattern matching operation in the UDWT domain. The complexity of the matching procedure can be reduced by at least an order of magnitude by using the UDWT local maxima to identify scales and delays that are most likely to maximize the detection statistic. The reduction can be achieved with little or no performance loss. The proposed detector is robust with respect to signal model and scale mismatch and discriminates against signals with scales outside the desired detection range.

7. REFERENCES

[1] B. Friedlander and B. Porat, "Performance analysis of transient detectors based on a class of linear data transforms," *IEEE Trans. Inform. Theory*, vol. 38, no. 2, pp. 665–673, Mar. 1992.

[2] M. Frisch and H. Messer, "Detection of a known transient signal of unknown scaling and arrival time," *IEEE Trans. Signal Processing*, vol. 42, no. 7, pp. 1859–1863, July 1994.

[3] S. Mallat and W.L. Hwang, "Singularity detection and processing with wavelets," *IEEE Trans. Inform. Theory*, vol. 38, no. 2, pp. 617–643, Mar 1992.

[4] P. Abry and P. Flandrin, "Multiresolution transient detection," in *Proc. of IEEE-SP Intl. Symposium on Time-Frequency and Time-Scale Analysis*, 1994, pp. 225–8.

[5] M. Frisch and H. Messer, "The use of the wavelet transform in the detection of an unknown transient signal," *IEEE Trans. Signal Processing*, vol. 38, no. 2, pp. 892–897, Mar. 1992.

[6] G. Wornell, *Signal Processing with Fractals: A Wavelet-Based Approach*, Prentice Hall, 1996.

[7] A.D. Whalen, *Detection of Signals in Noise*, Academic Press, 1971.

[8] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, 1992.

[9] T.T. Liu and A.C. Fraser-Smith, "Detection of transients in $1/f$ noise with the undecimated discrete wavelet transform," Manuscript in preparation, 1998.

[10] M. Alexandrescu, D. Gibert, G. Hulot, J. Le Mouël, and G. Saracco, "Detection of geomagnetic jerks using wavelet analysis," *J. Geophys. Res.*, vol. 100, no. B7, pp. 12557–12572, July 10, 1995.

[11] G. Corsini and R. Saletti, "Design of a digital $1/f^\gamma$ noise simulator," in *Noise in Physical Systems and $1/f$ Noise*, pp. 82–86. World Scientific, 1987.