POWER-LAW PROCESSORS FOR DETECTING UNKNOWN SIGNALS IN COLORED NOISE*

Ivars P. Kirsteins *Sanjay K. Mehta* *John Fay*

Naval Undersea Warfare Center Division Newport, Code 3124, Newport, RI 02841-5047 USA

ABSTRACT

We propose a new non-parametric adaptive detector for detecting an unknown broadband signal in interference consisting of non-stationary narrowband components and a locally stationary broadband component. An important feature of this detector is that it needs no prior information about the signal or interference. The proposed detector is based on the integration of the non-parametric power law detector of Nuttall with robust narrowband interference removal and whitening using a multiple taper spectral estimation-based technique. Experimental results indicate that the proposed detector outperforms conventional detectors.

1. INTRODUCTION

Our problem is the detection of a broadband signal in interference consisting of highly non-stationary narrowband components, a locally stationary colored broadband component, plus ambient noise as illustrated in figures 1 and 2. By locally stationary, we mean over some small time interval. The interference spectrum shown in figure 1 has the

Figure 1. Typical interference spectrum.

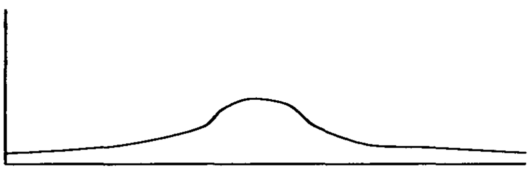

Figure 2. Typical signal spectrum.

form of a *mixed spectrum*, that is, consisting of a continuous spectrum component (broadband component and ambient noise) plus a discrete component due to tonals. Usually the signal and background noise spectrum and their respective probability density functions (pdf's) are unknown. Because the signal and noise pdf's and spectrum are unknown, it is not possible to find an optimum detector (e.g., Neyman-Pearson or uniformly most powerful invariant [1]). Standard adaptive methods [2] are also difficult to use since they generally require a specific model for the noise, say autoregressive, and similarily for the signal, of which we know nothing about. Furthermore, automated implementation of model-based adaptive detectors is not easy because the model type has to be determined and additional parameters such as the model order must be set on line. Since little is known about the noise and signal, we prefer a non-parametric adaptive detector, as opposed to a parametric detector.

2. POWER LAW DETECTOR

Recently, Nuttall [3] considered the problem of detecting a Gaussian signal in Gaussian noise, that if present, occupies an arbitrary set of M out of a total of N DFT bins, where M, the signal bin locations, and signal strength are all unknown. This is analagous to detecting a signal whose spectral shape is unknown. He proposed the following class of non-parametric *power law* processors

$$\sum_{k=1}^{N} \left(|X_k|^2\right)^v \underset{\text{Signal absent}}{\overset{\text{Signal present}}{\gtrless}} \lambda \tag{1}$$

where X_k is the kth data DFT bin

$$X_k = \int_\tau x(t) e^{-i2\pi \frac{(k-1)}{\tau} t}\, dt \quad , \tag{2}$$

$x(t)$ is the received time series, λ is a threshold, and v, the power, is a positive real number. Nuttall has determined experimentally that $v = 2.5$ yields best performance. Note that the test statistic in (1) is just a summation of the periodogram bins raised to the vth power. Note also that (1) does not use any information about the signal.

Nuttall derived (1) as a simple power approximation to the computationally impractical optimum likelihood ratio test (LRT) statistic (assuming M is known)

$$LRT \propto \sum_{k=1}^{K} e^{\alpha \mathcal{X}_k} \tag{3}$$

*THIS WORK WAS SPONSORED BY THE OFFICE OF NAVAL RESEARCH, MECHANICS AND ENERGY CONVERSION SCIENCE AND TECHNOLOGY DIV.

when the signal DFT bins have the same power, where α is some weight and

$$\mathcal{X}_k = \sum_{n \in L_k} X_n \quad , \text{ for } k = 1, 2, \ldots, K = \begin{pmatrix} N \\ M \end{pmatrix} \qquad (4)$$

is the summation over all possible sets L_k of M out of N bins. Since the number of permutations K is exceedingly large for most practical problems, the LRT is not computationally practical.

Remarkably, when v is appropriately chosen, the power law detector significantly outperforms the standard energy detector

$$\sum_{k=1}^{N} |X_k|^2 \quad \underset{\textit{Signal absent}}{\overset{\textit{Signal present}}{\gtrless}} \quad \lambda \qquad (5)$$

when the normalized signal bandwidth $M/N << 1$ [3]. In

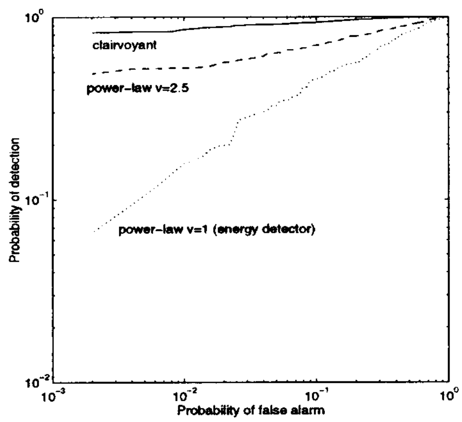

Figure 3. Computer simulation comparing power law detectors against clarivoyant detector for $M = 3$, $N = 200$, and $SNR = 12dB$.

figure 3 we show a simulation example comparing the optimum clairvoyant detector (when the signal bin locations are known) against the power law detector $v = 2.5$ and the energy detector $v = 1$ for $M = 3$, $N = 200$ and $SNR = 12db$. Here the power law detector significantly outperforms the energy detector and is close to the optimum detector. When signal bandwith is large, the power law detector performs nearly as well as the energy detector [3] (the energy detector is optimum when the signal occupies the entire Nyquist bandwidth).

The power law detector is attractive because it is non-parametric and needs no prior information about the signal spectrum and signal model, unlike AR methods. However, the power law detector assumes that $n(t)$ is white, which generally is not true in most practical applications where the interference components can be highly colored. Observing that the optimum detector for a signal in colored Gaussian noise consists of a pre-whitener followed by matched filtering [1], we propose the following adaptive power law detector:

$$\sum_k \left(\frac{\tilde{C}_2(f_k)}{\tilde{C}_1(f_k)} \right)^v \quad \underset{\textit{Signal absent}}{\overset{\textit{Signal present}}{\gtrless}} \quad \lambda \qquad (6)$$

where $\tilde{C}_1(f_k)$ is an estimate of the noise spectrum obtained from a signal-free *training* portion of data $\mathbf{x}_{\mathcal{T}} = [x_n x_{n+1} \cdots x_{n+P-1}]^T$ and $\tilde{C}_2(f_k)$ is an estimate of the observed data spectrum from the *detection* portion of the data $\mathbf{x}_{\mathcal{D}} = [x_{n+S} x_{n+S+1} \cdots x_{n+S+P-1}]^T$. The detector (6) structure also suggests that it is has constant false alarm rate (CFAR) properties, e.g., invariance to scalings. This is highly desirable in practical applications for calculation of thresholds and probability of false alarm.

Difficulties

However, it is extremely difficult to estimate the power spectrum with conventional periodogram-based methods (e.g. Bartlett or Welch Periodogram [4, 5]) when interfering tonal components are present, i.e., the data has a mixed spectrum consisting of a continuous and discrete part. Some of the difficulties are:

1) Tonal componenents tend to be highly non-stationary due to channel variability and source and receiver motion. Thus, tonal amplitude or power is difficult to estimate.
2) Over resolution of the spectral microstructure, which is inherently unstable due to non-stationarity, e.g., frequency drifting due to source and receiver motion, results in inaccurate estimates of the spectrum.
3) Unpredictable *smearing*, *biases*, and *leakage* effects due to windowing are inherent in periodogram-based methods.

Recall that the underlying broadband noise component with continuous spectrum was assumed to be locally stationary for the time encompassing the processing interval, but that the tonal components can be non-stationary. We propose to adaptively *separate* the non-stationary tonal components from the locally stationary continuous spectrum or broadband component. The separation is performed using the robust spectral estimation techniques of Thomson [4, 5], in which the undesirable tonal components are first removed or *cleaned* to obtain robust estimates of the background noise and the signal continuous spectrum components. These *cleaned* spectral estimates are then used to implement the adaptive power law detector.

In the next section we review multiple taper spectral estimation and the adaptive separation of the continuous part of the spectrum from the tonals. This is followed by a real data example comparing the performance of the proposed adaptive power law dector against the energy detector.

3. SEPARATION OF CONTINUOUS AND DISCRETE SPECTRUM

The separation of the continuous spectrum component from interfering tonals is not easy. Automatic removal of tonals using least-squares fitting can be difficult if the background spectrum has variations that are on the same order as the tonal levels (e.g., if a weak tonal occurs in a spectral valley, then the direct fitting procedure might *lock* onto a nearby peak of the background spectrum, rather than the tonal). Periodogram-based methods using smoothers, e.g., median, are also not robust and work poorly in the presence of closely spaced tones and high sidelobe leakage [4, 5].

3.1. Multiple Taper Spectral Estimation

The original goal of Thomson's multiple taper estimator [4, 5] was the spectral analysis of complicated non-stationary data consisting of lines plus a background component with continuous spectrum in which the physical processes generating the data were poorly understood. We now review the multiple taper method.

If we observe N contiguous samples $x_0, x_2, \ldots, x_{N-1}$ from a zero-mean stationary process, the Fourier transform of x_n is

$$\tilde{X}(f) = \sum_{n=0}^{N-1} x_n e^{-i2\pi fn} \qquad (7)$$

$X(f)$ can also be expressed as the convolution

$$\tilde{X}(f) = \int_{-1/2}^{1/2} \frac{sinN\pi(f-v)}{sin\pi(f-v)} dZ(v) \qquad (8)$$

where $dZ(f)$ is a zero-mean orthogonal process, with the true power spectrum $X(f)$ given by the expectation $X(f) = \mathcal{E}\{|dZ(f)|^2\}$ [4, 5].

Thomson [4, 5] proposed an approximate local solution of (8) for $dZ(f)$ in the frequency band $[f_0 - W, f_0 + W]$ in terms of the eigenfunctions of $sinN\pi(f-v)/sin\pi(f-v)$. One multiple taper spectral estimate [4, 5] is

$$\tilde{X}(f) = \frac{1}{K} \|V^H D(f)\mathbf{x}\|_F^2 \qquad (9)$$

where V is a $N \times K$ matrix whose columns are the principal *Discrete Prolate Spheroidal Sequence's* (DPSS's) $v_n^k(N, W)$, which are the eigenvectors of the $N \times N$ matrix

$$[R]_{m,n} = \frac{sin2\pi W(n-m)}{\pi(n-m)} \quad , \qquad (10)$$

arranged to correspond to the eigenvalues in descending order, $\mathbf{x} = [x_0, x_1, \ldots, x_{N-1}]^T$, $D(f) = diag(1, e^{-i2\pi f}, \ldots, e^{-i(N-1)2\pi f})$, and $K \approx 2NW$. W is chosen to be sufficiently small so that the true spectrum is approximately flat in the interval $[f_0 - W, f_0 + W]$, but large enough to keep the degrees of freedom $K = 2WN$ of $\tilde{X}(f)$ as large as possible.

Observing that V^H in (9) acts as an lowpass filter in the band $[-W, W]$, the projection of the frequency downshifted data vector $D(f)\mathbf{x}$ onto V^H becomes a bandpass filtering of the data to $[f-W, f+W]$ [4, 5]. We can now see that $\tilde{X}(f)$ is approximately the average energy in the band $[f - W, f + W]$. In other words, for a given set of frequency points, say $f_1, \ldots, f_M$, the spectrum estimate (9) is analogous to filtering $\mathbf{x}$ into the subbands $\{[f_0 - W, f_0 + W], [f_1 - W, f_1 + W], \cdots, [f_M - W, f_M + W]\}$, as shown in figure 4, and then calculating the average energy in each band.

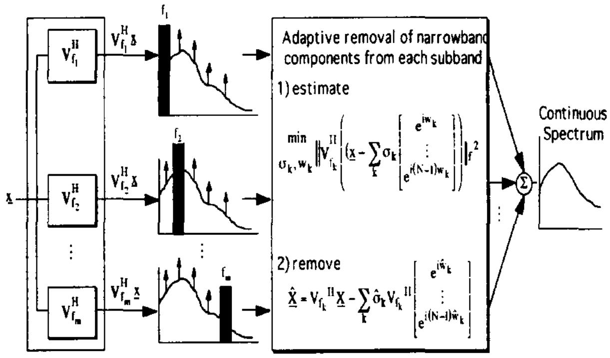

Figure 4. Depiction of multiple taper method and tonal removal.

3.2. Removal of Tonals

The multiple taper method provides a simple and effective way of locating and removing tonals from the underlying continuous spectrum background component. To mitigate the effects of background spectrum variations and adajent tonals, Thomson [4, 5] proposed that the estimation and removal of tonals be done within each subband. The key idea is that the projection of $D(f)\mathbf{x}$ onto V^H effectively isolates the frequency band $[f - W, f + W]$ from out of band tonals, and if W is properly chosen, the background noise spectrum is approximately locally *flat* or *white*. Thus, effects from out of band tonals and variations in the background spectrum are minimized.

To estimate and remove the tonals, Thomson used a simple least-squares fitting procedure which assumed that only one tonal could be present in the interval $[f - W, f + W]$. The procedure, is depicted in figure 4 (here generalized to an arbitrary number of tones). The single tonal assumption is a major restriction. In real data, one can have several tonals within each subband. An additional difficulty is that the tonals are generally non-stationary.

We now make the following modifications: To deal with non-stationarity, we partition the *training* $\mathbf{x}_{\mathcal{T}}$ and *detection* $\mathbf{x}_{\mathcal{D}}$ vectors of length N into L contigous smaller sub-blocks $\mathbf{x}_{\mathcal{T}}^T = [\mathbf{x}_{1,1}^T \mathbf{x}_{1,2}^T \cdots \mathbf{x}_{1,L}^T]$ and $\mathbf{x}_{\mathcal{D}}^T = [\mathbf{x}_{2,1}^T \mathbf{x}_{2,2}^T \cdots \mathbf{x}_{2,L}^T]$, such that the narrowband components are locally stationary within each subblock $\mathbf{x}_{j,k}$. The multiple tonals are then estimated in each subband using the method of Umesh [6] to solve

$$\min_{\alpha^f_{j,k,l}, w^f_{j,k,l}} \|V^H D(f)\mathbf{x}_{j,k} - \sum_{l=1}^{M_k^f} \alpha^f_{j,k,l} V^H \mathbf{e}(w^f_{j,k,l})\|_F^2$$
$$for\ j = 1,2\ ;\ k = 1,2,\ldots,L\ ;\ l = 1,2,\ldots,M_k^f \qquad (11)$$

(essentially, the complicated multidimensional least-squares optimization is solved through a sequence of simple 1-dimensional searches) and then removed, yielding the estimate of the continuous part of the spectrum at frequency f:

$$\hat{C}_j(f) = \frac{1}{L} \sum_{k=1}^{L} \frac{1}{K - M_k^f} \cdot \|V^H D(f)\mathbf{x}_{j,k} - \sum_{l=1}^{M_k^f} \hat{\alpha}^f_{j,k,l} V^H \mathbf{e}(\hat{w}^f_{j,k,l})\|_F^2 \qquad (12)$$

where the vector $\mathbf{x}_j$ corresponds to a data subblock from either the *training* or *detection* intervals and the $\hat{\alpha}^f_{j,k,l}$ and $\hat{w}^f_{j,k,l}$ are the solution to (11) in subband frequency f and subblock k. Note that $\mathrm{e}(w) = [1\, e^{iw}\, e^{i2w} \cdots e^{i(N-1)}]^T$ and the frequency search is restricted to the interval $[-W, W]$.

4. ADAPTIVE POWER LAW DETECTOR

Using the *cleaned* continous noise spectrum estimates derived above, the adaptive power law detector in (6) is implemented as follows:

1. Partition the data into *training* (noise only) and *detection* (noise plus signal) regions. Break up each region into smaller subblocks such that the narrowband interference is locally stationary.
2. Estimate the continuous part of spectrum $\tilde{C}_1(f)$ and $\tilde{C}_2(f)$ by removing the tonals using the steps outlined above (formulas (11) and (12)).
3. Form the test statistic

$$\sum_k \left(\frac{\tilde{C}_2(f_k)}{\tilde{C}_1(f_k)} \right)^v \underset{Signal\ absent}{\overset{Signal\ present}{\gtrless}} \lambda \qquad (13)$$

where $\tilde{C}_1(f)$ and $\tilde{C}_2(f)$ correspond to the estimates of the continuous spectrum in the *training* and *detection* regions respectively. The power v is determined empirically. We have found by simulation that $v = 2.5$ appears give best results over a wide range of conditions.

5. EXPERIMENTAL RESULTS

We now present some experimental results based on 240 seconds of real single channel noise time series collected in the ocean. The noise is characterized by the presence of numerous non-stationary interfering narrowband components from shipping, as shown by the spectrogram in figure 5. The proposed adaptive power detector was then applied to

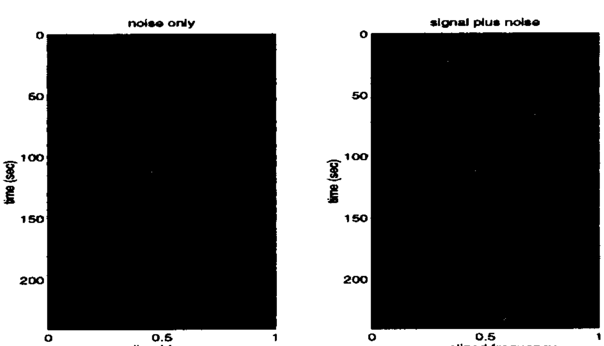

Figure 5. Spectrogram of real data with and without signal.

the real data with and without a simulated .1 hz passband Gaussian signal injected with $SNR = 12dB$. The signal plus noise spectrogram is shown in figure 5.

The detector parameters used were: $W = .055$ hz, adjacent 4.5 second blocks of *training* and *detection* data, $L = 10$ (subblock size of .45 seconds), $v = 2.5$, and the number of tonals determined by inspection. The measured receiver operating characterstic curves are plotted in figure 6 (based on a total of 105 trials). The curves show that the adaptive

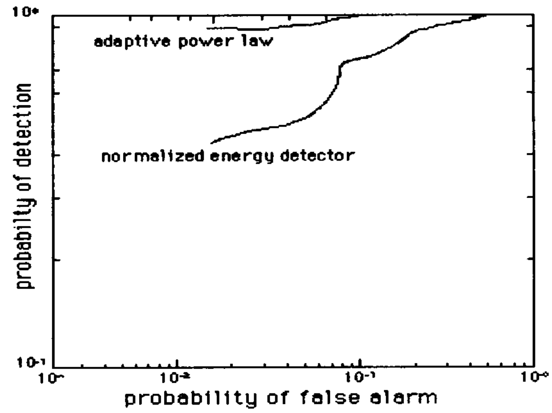

Figure 6. Adaptive power law detector compared against the normalized energy detector using real data and simulated signal.

power law detector performs much better than the normalized energy detector (normalized by an estimate of the noise variance in the *training* region).

Although not plotted here, we did try applying the power law detector to this data without any tonal removal or pre-whitening. As expected, it was sensitive to interfering tonals and performed poorly (since the power law detector is designed to detect tonals), thus verifying the need for robust interference removal and whitening.

6. CONCLUSION

We have presented a new adaptive non-parametric detector which needs little prior information about the signal, such as a detailed model or parameter values. A theoretical performance analysis of the proposed detector is needed in order to qualitatively predict detector performance and rate of adaptation. This will be done in future work.

REFERENCES

[1] H.V. Poor, *An introduction to signal detection*, Springer-Verlag, New York, 1988.

[2] S. Haykin, *Adaptive filter theory, third edition*, Prentice Hall, NJ, 1996.

[3] A.H. Nuttall, *Detection performance of power-law processors for random signals of unknown location, structure, extent, and strength*, NUWCDIVNPT Technical Report 10,751, 16 September, 1994.

[4] D.J. Thomson, "Spectrum estimation and harmonic Analysis," Proc. of IEEE, vol. 70, no. 9, pp. 1055-1096, Sept. 1982.

[5] D.J. Thomson, "An overview of multiple-window and quadratic-inverse spectrum estimation method," ICASSP-94, Adelaid, South Australia, vol. VI, pp. 185-194, April 1994.

[6] S. Umesh, *Fast maximum likelihood estimation of parameters in crowded signal environments*, Ph.D. dissertation, Univ. of Rhode Island, Kingston, RI, Aug. 1993.