AN APPROACH TO ADAPTIVE TRANSIENT DETECTION

Dragana D. Carevic

Maritime Operations Division, DSTO, Bldg A-51, HMAS Stirling, Rockingham, WA 6958, Australia

Abstract

This paper describes a detection method that adapts to unknown characteristics of the underlying transient signal, such as location, length, and time-frequency content. It applies a set of embedded detectors tuned to a number of signal partitions. The detectors are based on the general wavelet theory whereby two different techniques are examined, (1) the local Fourier transform (LFT) and (2) the discrete wavelet transform (DWT). The detection statistics are computed so as to enable prewhitening of unknown coloured noise and to allow for a constant false-alarm rate (CFAR) detection. An adapted segmentation of the signal is next obtained with a goal of finding the largest detection statistics within each segment of the partition. The detectors are tested using several underwater acoustic transient signals buried in ambient sea noise.

Introduction

Early detection in underwater acoustics is usually accomplished by using short-duration acoustic emissions, or transients, inadvertently emitted by the target. Typical duration of a transient ranges from few tens of milliseconds to several seconds. Signals are embedded in additive noise and the goal of the detection process is to determine whether the observations belong to a stationary noise distribution or not. For the detection problem treated in this paper it is assumed that the shape, length and location of the signals are not known.

Reference [1] describes several transient detectors that use the discrete Fourier transform (DFT) and the discrete wavelet transform (DWT) [2]. These detectors enable prewhitening of unknown correlated (coloured) noise and allow for a constant false-alarm rate (CFAR) detection. These properties are desirable for the passive sonar detection where the background noise is usually correlated and its statistics may change over time. However, a shortcoming of these detectors is that they use windowed signals where the length of the analysing window is fixed. These techniques perform well when the length of the analysing window is matched to the length of the transient, but when the window length is large as compared to that of the signal the performance of the detector deteriorates. By contrast, it is assumed that detection performance can be considerably improved by segmenting the signal into parts and by using the detection statistics related only to those segments that are characterised by large transform coefficients. The detection process in this way adjusts to the local characteristics of the signal.

Adaptive Transient Detectors

Similarly as in [1] the incoming data stream is divided into blocks of length N. The L+1 consecutive data blocks are considered where the (L+1)th block is the current block. The L previous data blocks are assumed to be noise-only and are used for background normalisation. The blocks are processed by using either the local Fourier transform (LFT) [3],[4] or the DWT [2], resulting in two somewhat different adaptive detection procedures.

The LFT-Based Detectors

The LFT provides a means for expanding a signal into a set of smooth orthonormal bases subordinate to an arbitrary partition (segmentation) of the signal. The bases consist of complex exponential function smoothly restricted to adjacent overlapping segments. The bases are 'local' in the sense that they expand the signal only within the particular region of interest. Starting with a partition of the signal into disjoint intervals $\mathbf{R} = \cup_k I_k = (\alpha_k, \alpha_{k+1})$ (see for example Figure 1) the method uses the smooth orthogonal periodisation [3]. It first performs windowing of the original signal using smooth overlapping compactly supported bump functions or windows placed over each interval. This is followed by folding the windowed segments to the disjoint intervals. Each interval is next locally periodised and separately processed using the standard discrete Fourier transform (DFT). The method ensures that the resulting basis set is orthonormal for $L^2(\mathbf{R})$. Also, the signal decomposition is non-redundant, that is the number of (complex) transform coefficients is the same as the length of the signal. Since the basis functions are smooth, no discontinuity is created at the segment endpoints.

The LFT is applied to a binary tree structured set of intervals yielding a redundant M-level signal decomposition (see Figure 2, also in [3]). The length of the segment at the lowest decomposition level (m=0) is dyadic $N=2^K$ and equals to the block-length of the analysed signal. The intervals at the higher decomposition levels m=1,2,…,M-1 are obtained by recursively splitting the segments at lower levels at the mid-point. In this way, at each level m=0,1,…,M-1 there are 2^m segments of length $2^{(K-m)}$. The maximum number of decomposition levels is M<=K and M is chosen with respect to the smallest length of the signal segments of interest. Note that the LFT coefficients can be taken to represent disjoint intervals, since the overlap of the windowed segments can be made relatively small (only a few samples).

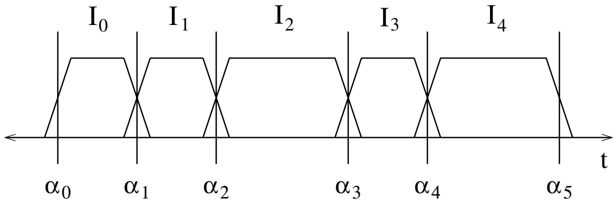

Figure 1. Partition of line and the corresponding windows for smooth localised orthonormal basis.

We compute detection statistics for each segment $q=1,2,\ldots,2^m$ at levels $m=0,1,\ldots, M-1$ within the binary tree. L previous data-blocks are used to estimate the background statistics for normalisation. These L blocks are also processed using the M-level LFT. Consider the q-th segment at the m-th level of decomposition that belongs to the i-th data block, $i=1,2,\ldots,L+1$. Then define $U^{(1)}_{m,q,j,i} = X_{m,q,j,i}$ and $U^{(2)}_{m,q,j,i} = X_{m,q,j-1,i}+X_{m,q,j,i}$ where $X_{m,q,j,i}$, $j =1,2,\ldots,2^{(K-m)}$ is the j-th magnitude squared transform coefficient within this segment. Two detection statistics $TF^{(n)}_{m,q}$ for n=1,2 are then defined by

$$TF^{(n)}_{m,q} = \sum_{j=1}^{2^{(K-m)}} \left(\frac{U^{(n)}_{m,q,j,L+1}}{\frac{1}{2^m L}\sum_{i=1}^{L}\sum_{p=1}^{2^{(K-m)}} U^{(n)}_{m,p,j,i}} \right)^{\nu} . \qquad (1)$$

At the level of decomposition m=0 there are L background segments that are used for normalisation. The statistics computed at this level, $TF^{(n)}_{0,1}$, n=1,2, are the same as those defined in reference [1] for the block-length 2^K. At the level m, $m=1,2,\ldots,M-1$, there are $2^m L$ background segments from L previous blocks, for $i=1,2,\ldots,L$ and for $q=1,2,\ldots, 2^m$. These segments are used to compute detection statistics in all 2^m segments at the m-th binary tree level of the current (L+1)-th block.

The normalisation of each component carried out in Eq. (1) prewhitens correlated noise and allows for CFAR detection performance. The detection statistics $TF^{(2)}_{m,q}$ is defined such as to exploit signal contiguity in the frequency domain. Namely, a transient signal is usually bandpass and it is reasonable to expect that most of its energy will be spread over a contiguous band in the frequency domain.

The exponent ν in Eq. (1) is an adjustable parameter.

The intention is to find those segments in the binary tree structure that have the largest values of the detection statistics, as this indicates that these segments are best matched by their basis functions. In order to be able to compare the values of $TF^{(n)}_{m,q}$ they need to be further normalised and this is done by using the means and variances of the $TF^{(n)}_{m,q}$ estimated in the noise-only case. The $TF^{(n)}_{m,q}$ for each segment are computed as sums of random variables (rv's) and in the noise-only case their distribution depends on the number of the rv's used in the summation (that is, on the length of the underlying segment). So under the assumption that noise has a stationary distribution the noise-only $TF^{(n)}_{m,q}$'s related to the segments at the decomposition level m have identical distributions, and their distribution varies across different levels. Using the central limit theorem Wang and Willet [1] assume that these detection statistics converge to normal distributions and provide formulae for calculating means and variances of these distributions. In our experiments using real sea noises it was found that these formulae do not estimate means and variances of the detection statistics across the different levels of distribution very accurately. The probable reason is that the normal distribution assumption is usually not correct and the formulae are computed using approximations anyway. Therefore, in this paper means and variance used for normalisation of the detection statistics $TF^{(n)}_{m,q}$, n=1,2, for all segments $q=1,2,\ldots,2^m$ at the levels $m=0,1,\ldots, M-1$ of the binary tree are estimated using a sufficiently long noise-only section of the input data stream.

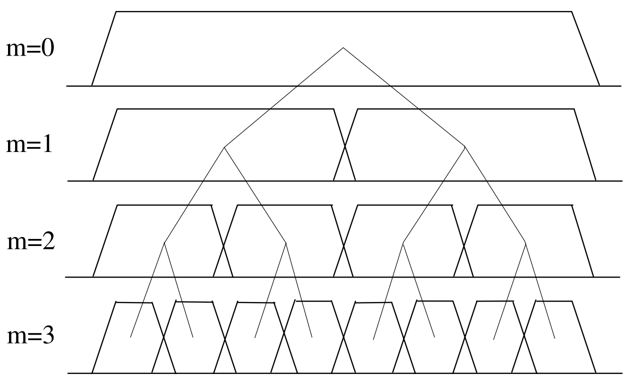

Figure 2. Organisation of localisation intervals into a binary tree.

The DWT-Based Detectors

We now describe a method for computing a nested binary tree-structured set of detectors based on the DWT. Since the incoming data stream is assumed to be very long, the DWT is defined as applied to a doubly infinite input sequence [3]. In particular, DWT is computed for separate data blocks where the initial conditions for processing the current block are obtained by carrying over the information resulting from processing the immediately preceding data block. Consider a full (K scale-levels) DWT of the block of length 2^K (high-pass coefficients) computed in this way. This DWT contains within itself the full DWT's of the entire binary tree structured set of intervals obtained by recursive mid-point splitting of the lower level segments. As an example, Figure 3 shows the DWT coefficients for the block-length 32. It can be seen that this DWT contains the DWT's of the 4 successive intervals of length 8 at the level m=2 of the binary tree structure, and the DWT's of the 2 intervals of length 16 at the level m=1. The coefficients of these DWT's are grouped as shown in the figure. Consequently, the DWT coefficients of the nested set of intervals within the binary tree can be obtained simply by rearranging the DWT coefficients computed at the lowest level of the structure, that is, for m=0.

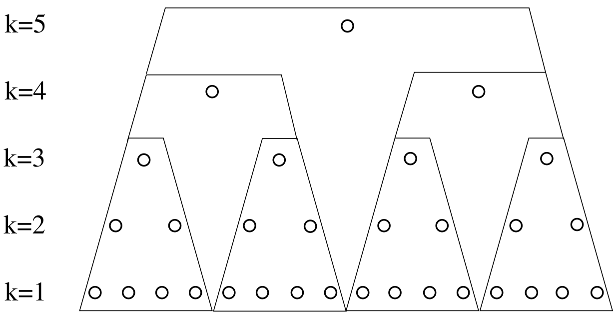

Figure 3. Organisation of the DWT coefficients of localisation intervals into a binary tree. Here *k* denotes the levels of the DWT decomposition.

We compute the detection statistics using the DWT coefficient within each segment $q=1,2,\ldots,2^m$ at levels $m=0,1,\ldots, M-1$ of the binary tree related to the current data block. It is assumed that we also have the DWT coefficients of the L immediately preceding noise-only data blocks. Consider the q-th segment at the m-th level of the i-th data-block, $i=1,2,\ldots, L+1$. Then let $Y_{m,q,k,j,i}$ denote the magnitude-squared DWT coefficient computed for the data within this segment, where $k=1,2,\ldots,K-m$ denotes the DWT scale index and $j = 1,2, \ldots, 2^{(K-k-m)}$ is the within-scale time index. Next define $U^{(1)}_{m,q,k,j,i} = Y_{m,q,k,j,i}$ and $U^{(2)}_{m,q,k,j,i} = Y_{m,q,k,j,i}+Y_{m,q,k-1,2j-1,i}+Y_{m,q,k-1,2j,i}$ for $k = P^{(n)}_m,\ldots,K-m$, $j = 1,2,\ldots,2^{(K-k-m)}$, and $P^{(1)}_m = 1$ and $P^{(2)}_m = 2$. The detection statistics $TW^{(n)}_{m,q}$ are then given by

$$TW^{(n)}_{m,q} = \sum_{k=P^{(n)}_m}^{K-m} \sum_{j=1}^{2^{(K-k-m)}} \left(\frac{U^{(n)}_{m,q,k,j,L+1}}{\frac{1}{2^m L}\sum_{i=1}^{L}\sum_{p=1}^{2^m} U^{(n)}_{m,p,k,j,i}} \right)^{\nu} \qquad (2)$$

for n=1,2. Here, similarly as in the previous in section, the statistics are computed using the normalised wavelet coefficients, and $TW^{(2)}_{m,q}$ is defined such as to exploit contiguity of transient signals in the spatial-frequency domain. The rationale for defining the statistics $TW^{(2)}_{m,q}$ is that a transient signal can be expected to have components that are large over a contiguous band in both time and frequency domains. Also, the detection statistics $TW^{(n)}_{m,q}$ n=1,2 computed at the level m=0 are equivalent to those defined reference [1].

All detection statistics computed in this way are again normalised for each decomposition level $m=0,1,\ldots,M-1$ so as to have zero mean and unit variance. This is done using the means and standard deviations estimated based on a noise-only section of the input data stream.

A disadvantage of the standard (critically sampled) DWT is that it is non-invariant to time-shifts of the analysed signal. The input-signal shifts can generate unpredictable changes in the DWT coefficients that can cause a degradation of the performance of the detectors. Several techniques using best-basis or optimal wavelet designs have been proposed to reduce DWT shift sensitivity [5]. An alternative approach is to use complex wavelets to compute the DWT, with the real and imaginary parts of the complex wavelet constituting a Hilbert transform pair. It is argued that, if a transient is present, the real and imaginary part of the complex transform coefficients cannot simultaneously be small and the magnitude of the coefficients should be used for detection [6]. Recently Fernandes [7] proposed a linear projection filter that projects a real-valued signal onto the 'Softy' space. The filter has the pass-band over over $[0,\pi]$ and stop-band over $[-\pi,0]$ so that it retains positive frequencies and suppresses negative frequencies. The standard DWT is applied to the projected (complex) input signal obtained in this way and as the result complex wavelet coefficients are generated [7]. This projection-based complex DWT is applied to compute the DWT coefficients in this section.

Adapted Signal Partition

The processing described in the previous sections results in a set of detectors associated with the binary tree structure of intervals $I_{m,q}$ in Figure 2. We seek the partition (segmentation) $\cup I_{mi,qi}$ of the current block restricted to $I_{m,q}$, such that $\cup I_{mi,qi}$ exhaustively covers the block-length 2^K without overlapping, and provides the largest values of the resulting detection statistics. Let $T_{m,q}$ denote the normalised detection statistics irrespective of the method used for its computation (*i.e.*, obtained using either Eq. (1) or Eq. (2)). Set $J_{m,q}$ to the most refined partition in the binary tree structure, $J_{M-1,q} = I_{M-1,q}$, $q=1,2,\ldots,2^{M-1}$, and, also, set $\theta_{m,q} = T_{m,q}$. Then the following recursive search algorithm

$$J = \begin{cases} I_{m-1,q} & \theta_{m-1,q} > \max(\theta_{m,2q-1}, \theta_{m,2q}) \\ J_{m,2q-1} \oplus J_{m,2q} & otherwise \\ \theta_{m-1,q} = \max(\theta_{m,2q-1}, \theta_{m,2q}) & \end{cases} \qquad (3)$$

yields the required partition $J_{0,1} = \cup I_{mi,qi}$. Let $\{T_{mi,qi}\}$ be the set of detection statistics associated with the partition $\cup I_{mi,qi}$. Then it can be shown that

$$\max_{m_i,q_i}(T_{m_i,q_i}) = \max_{m,q}(T_{m,q})$$

that is, the segment for which the detection statistics is maximal over the entire binary tree structure is included in the best signal partition obtained using Eq. (3).

We next assign the values from the set $T_{mi,qi}$ to the segments that correspond to the finest partition of the binary tree $I_{M-1,q}$, $q=1,2,\ldots,2^{M-1}$ as follows. The segments from the partition obtained using Eq. (3) that belong to the set $I_{M-1,q}$, that is for which m_i=M-1, are assigned their respective value $T_{mi,qi}$. Each segment $I_{mi,qi}$ for which m_i<M-1 is recursively split into halves until the highest level (M-1) is reached. To all intervals obtained in this way the same value of the detection statistics is assigned that is equal to the value corresponding to their parent segment. As the result a temporal sequence of detection statistics of the length $2^{(M-1)}$ is obtained. It characterises the underlying data block and is used for transient detection. Namely, each value of the detection statistics from the sequence, t, is separately compared to a threshold τ. The probability that t exceeds the threshold τ for the noise-only data (hypothesis H_0) defines the probability of false alarm P_{fa} whereas the probability that $t>\tau$ when a transient signal is present in the data (hypothesis H_1) is denoted as the probability of detection P_d.

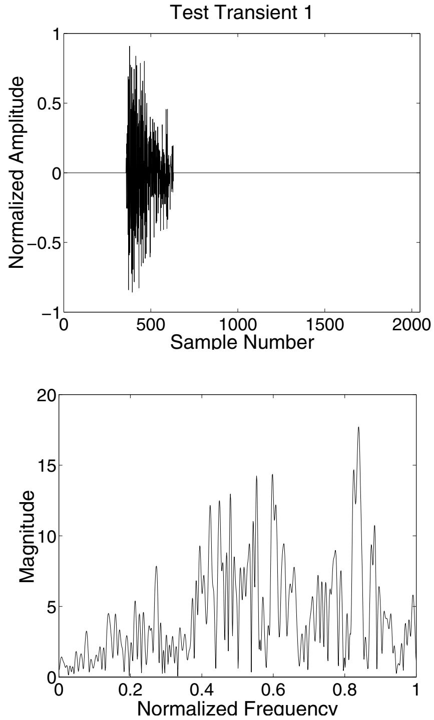

Figure 4. Test transient 1 used in the experiment. Time series representation (top), frequency domain representation (bottom).

Detector Performance

The performance of the LFT- and DWT-based detectors is evaluated using two underwater acoustic transients (see Figures 4 and 5). As can be seen the transients vary in their characteristics regarding both duration and frequency content. A 40 minute long recording of ambient sea noise is used as the background signal. This signal is divided into blocks of length N=2048 samples and the transients are inserted into these blocks at different signal-to-noise ratios (SNR's). This is followed by computing the detection statistics using an M=5 level binary tree decomposition. It is assumed that, in some cases, smaller parts of a transient signal may be better matched by the transform basis functions than the entire signal, and that detection performance can be improved by using these segments. Therefore, the length of the smallest segment in the binary tree, ie., the one that corresponds to the highest decomposition level M-1, is set to 128 samples, which is smaller than the length of any of the tested transients. The starting position of a transient within a block is chosen such as not to coincide with the beginning of any subinterval at any level of the binary tree structure and to allow that the entire transient is contained within the block.

The number of noise-only data blocks used in Eqs. (1) and (2) for background normalisation is set to L=10.

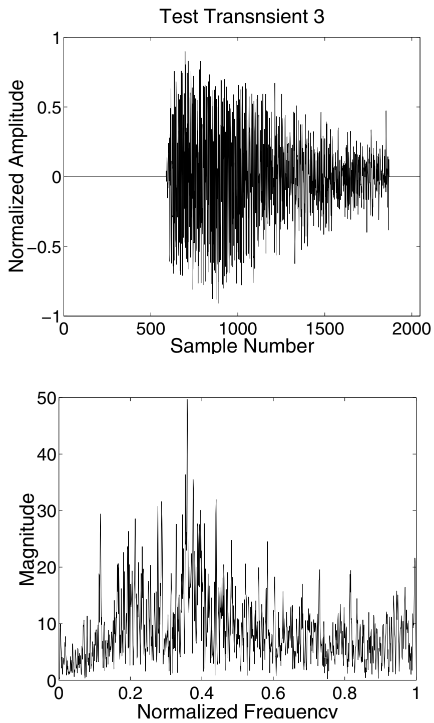

Figure 5. Test transient 3 used in the experiment. Time series representation (top), frequency domain representation (bottom).

Regarding the choice of the exponent ν in Eqs. (1) and (2) the study in [1] indicates that, for different block-lengths N and output signal-to-noise ratios (SNR's), and under the assumption that the frequency characteristics of the transient are not known, good values are $1.5<\nu<2$. In our experiments we obtain similar results for both the LFT- and the DWT-based detectors. We therefore set $\nu=1.7$. Also, we tested several different wavelets with different lengths of the associated filters and found that the performance of the DWT-based detectors does not depend much on the choice of wavelet. We use the Haar wavelet. The reason for this choice is that the filter that corresponds to the Haar wavelet is short, so that the information carried over from one block of data to the other in the processing of the on-line DWT is minimal.

In the experiments the total number of data blocks of length N=2048 of the sea noise record is B=9600. For both LFT- and DWT-based detectors we use the statistics of the first 0.15B noise-only blocks to estimate means and variances of the detection-statistics over different decomposition levels of the binary tree. These estimates are next used to normalise the detection statistics computed for the remaining 0.85B data blocks.

Probability of detection for the LFT- and DWT-based detectors. In the experiments we have only one occurrence of a transient in each data block. For this reason a transient is considered detected if any of the detection statistics from the temporal sequence obtained using the best-adapted partition of the signal, restricted to the segments in which resides 90% of the transient energy, is greater than a given threshold. The sequence. The probability of detection is defined by

$$P_d = \frac{B_d}{B_{tot}}$$

where B_d is the number of blocks in which a transient is detected and B_{tot} is the total number of data blocks.

Probability of false alarm for the LFT- and DWT-based detectors. False alarm is assumed to occur each time a detection statistics from the sequence resulting from the best-adapted partition of a noise-only data block is larger than the threshold. Therefore, the probability of false alarm is defined as

$$P_{fa} = \frac{\sum_{b=1}^{B_{tot}} N_b}{2^{M-1} B_{tot}}$$

where N_b is the number of values in this sequence that are larger than the threshold for the block $b=1,2,\ldots,B_{tot}$, and M is the number of levels of the binary tree structure of intervals.

We evaluated the performance of the LFT- and DWT-based detectors for the transients used in this study and compared it to the performance of the DFT- and DWT-based detectors that use fixed window length described in [1]. The performance of the DFT- and DWT-based detectors in [1] is evaluated for the window length N=2048. The corresponding detection statistics $TF^{(n)}_{0,1}$ and $TW^{(n)}_{0,1}$ for n=1,2 are computed using the transform coefficients at the level m=0 of the binary tree structure as per Eqs. (1) and (2) respectively.

Figures 6 and 7 show receiver operating characteristics (ROC) evaluated for a fixed SNR specific to each test transient where the detectors are denoted as follows. The LFT-based detectors are denoted by LFT-T1 for the detection statistics $TF^{(1)}_{m,q}$, and by LFT-T2 for $TF^{(2)}_{m,q}$ in Eq. (1), whereas the DWT-based detectors are denoted by LWT-T1 and LWT-T2 for $TW^{(1)}_{m,q}$ and $TW^{(2)}_{m,q}$ in Eq. (2), respectively. The corresponding DFT-based detectors described in [1] are denoted by DFT-T1 and DFT-T2 for $TF^{(1)}_{0,1}$ and $TF^{(2)}_{0,1}$, and the DWT-based detectors by DWT-T1 and DWT-T2 for $TW^{(1)}_{0,1}$ and $TW^{(2)}_{0,1}$, respectively.

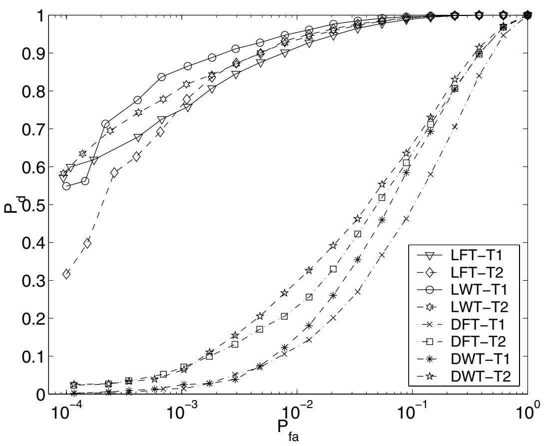

Figure 6. Receiver operating characteristics for the test transient 1 at the SNR fixed at –5 dB.

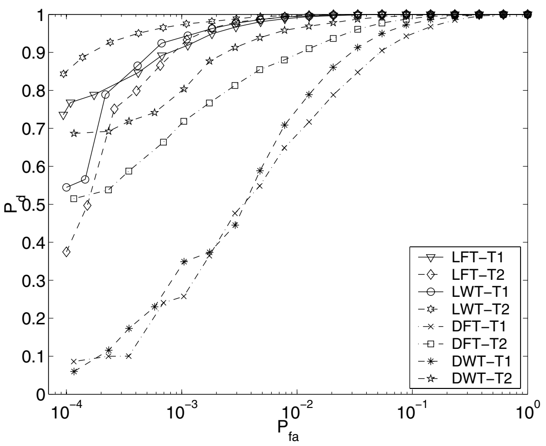

Figure 7. Receiver operating characteristics for the test transient 3 at the SNR fixed at –8 dB.

The results in Figures 6 and 7 indicate that adaptive window-length detectors perform better than the detectors described in [1] for the tested transients. Also, the adaptive window-length detectors using the DWT outperform the LFT-based detectors. For the test transient 1 the LWT-T1 detector gives somewhat better result that the one obtained using the detector LWT-T2, whereas, for the test transient 3, the detector LWT-T2 performs better than the detector LWT-T1.

Conclusions

This paper is concerned with the detection of acoustic underwater transients of unknown location, length, and time-frequency content. A method that applies a set of embedded transient detectors tuned to a number of signal partitions has been proposed. The detectors are based on the general wavelet theory whereby two different techniques are examined, the local Fourier transform and the discrete wavelet transform. The detection statistics are computed so as to enable prewhitening of unknown coloured noise and to allow for a constant false-alarm detection rate. The statistics are combined so as to obtain a best-adapted partition of the signal with the goal of finding the largest detection statistics in each segment.

The detectors are tested using several underwater acoustic transients buried in ambient sea noise. The results show that the proposed adaptive window-length detectors outperform the detectors with fixed window lengths.

References

[1] Wang Z, Willett P, "All-purpose and plug-in power-law detectors for transient signals," *IEEE Trans. Signal Process.*, 49(11):2454-2466, 2001.

[2] Mallat S, "A theory of multiresolution signal decomposition: the wavelet representation," *IEEE Trans. Pattern Anal. Machine Intell.*, 11:674-693, 1989.

[3] Wickerhauser M V, *Adapted wavelet analysis from theory to software*, Pitscataway, NJ: IEEE Press, 1994.

[4] Saito N, "Local Fourier dictionary: a natural tool for data analysis," in Proc. SPIE, Wavelet Application in Signal and Image Processing, 3813:610-624, 1999.

[5] Liang J, Parks T W, "A translation-invariant wavelet representation algorithm with applications," *IEEE Trans. Signal Process.*, 44(2):225-232, 1996.

[6] Abry P, Flandrin P, "Multiresolution transient detection,'" in Proc. of the IEEE-SP Intern. Symp. Time-Frequency and Time-Scale Analysis, 225-228, Philadelphia, PA, USA, 1994.

[7] Fernandes F C, *Directional, shift-insensitive, complex wavelet transforms with controllable redundancy*, PhD thesis, Rice University, Houston, Texas, USA, 2002.