DETECTION OF GAUSSIAN BANDPASS TRANSIENTS UNDER IMPULSIVE NOISE: A WAVELET TRANSFORM APPROACH

Francisco M. Garcia *Isabel M. G. Lourtie*

ISR - Instituto de Sistemas e Robótica
IST - Instituto Superior Técnico
Torre Norte, Av.Rovisco Pais
P-1096 Lisboa Codex Portugal
E-mail : fmg@isr.isr.ist.utl.pt

ABSTRACT

In underwater acoustics, the modeling of impulsive noise ambients by symmetric-α-stable laws is motivated by the generalized central limit theorem. However, detection of stochastic signals under such additive noise is a difficult task to implement, due to the lack of a closed-form expression of the *a-posteriori* probability density function. In this paper, we present a suboptimal detector for Gaussian bandpass transients in impulsive noise that uses a nonlinear, memoryless prefilter followed by a discrete wavelet transform. The resulting signals present a Gaussian-like behavior and the decision is achieved by the comparison of a quadratic likelihood ratio with a threshold. The tuning of the nonlinearity parameter is performed either by looking at the receiver operating characteristic or using the Chernoff distance, that, although resulting in an approximate solution, is easier to compute. Simulation results are presented by Monte-Carlo simulation.

1. INTRODUCTION

In many signal processing applications, the assumption of Gaussianity for the ambient noise is not realistic. For example, in underwater acoustics, the effects of ice-cracking, and some forms of reverberation and rifting are frequently referred [1, 2] as sources of impulsive noise. This is characterized by the presence of occasional bursts or sharp spikes, corresponding to heavier probability density function (pdf) tails when compared to Gaussian distributions. This behavior frequently results in distributions with infinite variance, stable processes being often used to model such phenomena.

Stable distributions, although lacking some of the nice results that are available in the Gaussian case for statistical signal processing, have the advantage of satisfying a generalized central limit theorem, that states that the sum of a large number of independent and identically distributed (i.i.d.) random variables with or without infinite variance converges to a stable law. The inexistence of a closed form expression for the pdf of a stable process, except in some particular cases (in general, stable laws are described by their characteristic function) turns difficult the development of an optimal likelihood processor.

Detection, classification and localization of signals in impulsive noise have been studied in the last years [1, 2, 3, 4]. Due to the difficulties in handling non-Gaussian distributions, these works assume that the noise is independent and the signals are deterministic. However, in passive detection and classification, this latest assumption is unrealistic.

In this paper, we present a suboptimal binary detector for Gaussian distributed discrete transient signals embedded in impulsive noise with infinite variance. The two hypotheses consist in noise only (H_0), and noise plus signal (H_1), respectively. The processor is represented by a series of four blocks (fig. 1): the observation process is first transformed through a nonlinearity that modifies the pdf of each hypothesis in such a way that its output has finite variance. Then, the observation is decomposed under the discrete wavelet transform (WT). At this point, and under both hypotheses, the observation process distribution is close to a Gaussian, depending on the scale at which the WT is performed and on the length of the wavelet filters. Based on the covariance matrices of the process at the output of the WT block under both hypotheses, the likelihood ratio is computed assuming Gaussianity, and compared to a threshold. An optimization procedure for the tuning of the nonlinearity is presented. Assuming that the ambient impulsive noise obeys a symmetric-α-stable (S-α-S) distribution, the receiver performance is addressed by Monte Carlo simulation.

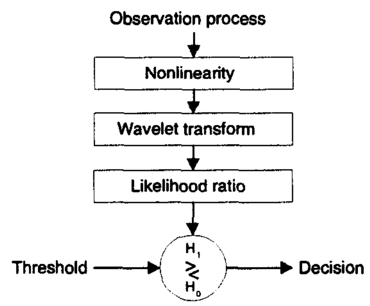

Figure 1. Processor block diagram.

2. PROBLEM FORMULATION AND DETECTOR STRUCTURE

Let us assume that the observation process $r(i)$ is such that

$$r(i) = \begin{cases} n(i), & \text{under hypothesis } H_0 \\ s(i) + n(i), & \text{under hypothesis } H_1, \end{cases}$$

where i is the discrete time variable, $n(i)$ is a white impulsive noise sequence with zero mean (or median), mutually independent of $s(i)$, which is a discrete-time zero mean

0-8186-7919-0/97 $10.00 © 1997 IEEE

Gaussian distributed bandpass transient with covariance matrix $A(i,j) = E[s(i)s(j)]$. Defining the discrete time-frequency power spectral density, $S(\Omega, i)$, as the discrete Fourier transform of the autocorrelation function $A(i, i+p)$ over the lag variable p, the process $s(i)$ exhibits a limited bandwidth in the frequency domain, i.e.,

$$\forall i \quad S(\Omega, i) = 0, \qquad \text{if} \qquad |\Omega| \in [0, \Omega_{min}[\cup]\Omega_{max}, 2\pi],$$

and a nearly finite duration in the time domain, i.e.,

$$\forall \Omega \quad S(\Omega, i) \simeq 0, \qquad \text{if} \qquad i \notin [0, N_s].$$

The nonlinear block is the hole puncher represented in fig. 2. The distribution of the process at the output of this block has a finite variance that depends on the parameter a. If a is too small or too large, then a poor performance of the processor is expected: the variances in both hypotheses will be either too small or too large and the separability between the two hypotheses is small. In the sequel, we present a procedure to tune the parameter a in order to minimize the Chernoff distance between the two hypotheses.

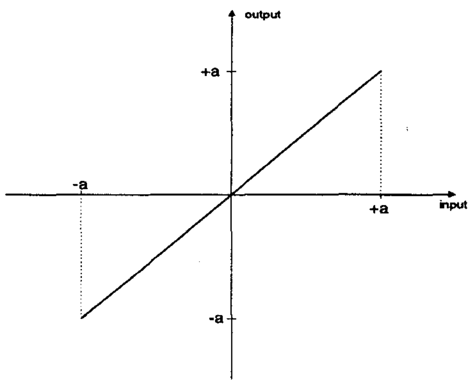

Figure 2. Hole puncher nonlinearity.

The second block of fig. 1 operates a wavelet transform at the output of the nonlinearity. In [5, 6], we showed that the WT is suited for the decomposition of Gaussian bandpass transients, by reducing their complexity, leading to sparse covariance matrices of small order, and being adapted to real-time processing.

Using Mallat's recursive algorithm for image decomposition [7, 8] in its one-dimensional form,

$$\begin{aligned} c_k^j &= \sum_i h(i-2k)c_i^{j-1} \\ d_k^j &= \sum_i g(i-2k)c_i^{j-1} \end{aligned}, \tag{1}$$

where $h(i)$ and $g(i)$ are Daubechies [8] finite length filters, the observed prefiltered discrete time sequence $c_i^0 = r(i)$ is decomposed in the subsequences d^1, d^2, ..., d^m and c^m. The recursive filtering equations (1) can also be expressed in terms of the original signal $r(i)$ as the internal product

$$\begin{aligned} c_k^j &= < r, h_k^j > \\ d_k^j &= < r, g_k^j > \end{aligned}. \tag{2}$$

Filters $h(i)$ and $g(i)$ are, respectively, lowpass and highpass filters. While c^m represents a smoothed version of r, sequences d^j stand for filtered versions of r in different frequency bands [5, 6]. The internal product $< r, g_k^j >$, is relevant only when both signal r and filter g_k^j exhibit overlapping frequency bands. Consequently, performing the WT decomposition of r until the lowpass residue corresponding to $< r, h_k^m >$ is close to zero, and neglecting those sequences d^j for which $< r, g_k^j >$ is also close to zero, one gets an approximate representation of r based on a smaller set of frequency scales. Thus, a single coefficient d_k^j is, under both hypotheses, the result of a weighted sum of random variables,

$$d_k^j = \sum_{i=1}^{N_j} r(i)z(i)g_k^j(i) \tag{3}$$

where N_j is the length of the filter g_k^j at scale j, and

$$z(i) = \begin{cases} 1, & \text{if } |r(i)| \leq a \\ 0, & \text{otherwise} \end{cases}$$

stands for the nonlinear function represented in fig. 2.

Under hypothesis H_0, the terms $r(i)z(i) = n(i)z(i)$ are i.i.d. with finite variance. By Ljapunov's central limit theorem (CLT), the infinite sum of independent random variables (RV), not necessarily with the same variance or distribution, tends to a Gaussian if and only if they satisfy the Lindberg condition [9]. Accordingly to this condition, a RV that is composed by a large number of independent infinitesimal RV (i.e., whose dispersions are very small with respect to the dispersion of the sum), has a distribution which, for all practical purposes, may be regarded as normal. This is the case for large values of both the filter length N_j and the nonlinearity parameter a (corresponding to a high probability of $z(i)$ taking the value 1). Since the filters g_k^j form an orthonormal set, the coefficients covariance matrix under hypothesis H_0 is

$$C_{H_0} = \sigma_a^2 I,$$

where I denotes the identity matrix, and

$$\sigma_a^2 = \int_{-a}^{a} x^2 f_n(x) dx,$$

where $f_n(x)$ is the pdf of the impulsive noise. Although it may not have a closed-form expression, in the case of stable noise it is easily computed by taking the inverse Fourier transform of the characteristic function.

Under hypothesis H_1, we have

$$d_k^j = \sum_{i=1}^{N_j} s(i)z(i)g_k^j(i) + \sum_{i=1}^{N_j} n(i)z(i)g_k^j(i). \tag{4}$$

As for hypothesis H_0, for large values of a, the second term on the right-hand side (rhs) of eq. (4) is approximately Gaussian with a covariance matrix given by C_{H_0}. Regarding the first term, let us point out that, for large a, the nonlinear function $z(i)$ can be approximated to

$$z(i) \simeq \begin{cases} 1, & \text{if } |n(i)| \leq a \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

In fact, although $z(i)$ is a random sequence depending on both $n(i)$, $s(i)$ and the parameter a, for values of a sufficiently large compared to the maximum standard deviation of the process, we can assume that the signal is, in general, not affected by the nonlinearity, except when the noise bursts largely emerge. In the above context, the first term on the rhs of eq. (4) consists mainly in a linear decomposition of a Gaussian process, being approximately Gaussian distributed. Consequently, the global pdf of d_k^j is also approximately Gaussian.

From (3), the elements of the coefficients covariance matrix C_{H_1} are given by

$$E[d_k^j d_p^q] = \sum_{i,l} E[z(i)z(l)r(i)r(l)g_k^j(i)g_p^q(l)]. \quad (6)$$

As referred to before, for a sufficiently large, $z(i)$ depends mostly on the impulsive noise, see eq. (5). Finally, since transient signal and noise are mutually independent, the elements of the covariance matrix C_{H_1} become

$$\begin{aligned} E[d_k^j d_p^q] &\simeq \sigma_a^2 \delta_{j-q}\delta_{k-p} + \\ &\quad + \sum_{i,l} A(i,l)E[z(i)z(l)]g_k^j(i)g_p^q(l) \\ &\simeq \sigma_a^2 \delta_{j-q}\delta_{k-p} + E[z^2]\Gamma 1 + E[z]^2\Gamma 2, \end{aligned} \quad (7)$$

where

$$\Gamma 1 = \sum_i A(i,i) g_k^j(i) g_p^q(i)$$

and

$$\Gamma 2 = \sum_i \sum_{l \neq i} A(i,l) g_k^j(i) g_p^q(l).$$

In (7), for the nonlinear function considered herein,

$$E[z] = E[z^2] = \int_{-a}^{a} f_n(x)dx.$$

The parameter a needs to be adapted in order to minimize the receiver probability of error (PE). This procedure is achieved by Monte Carlo simulation, and by looking at the receiver operating characteristics (ROCs). For every value of a, it is necessary *i)* to compute the covariance matrices at the output of the WT block under both hypotheses, and *ii)* to proceed with Monte Carlo simulation in order to determine the ROCs. Alternatively, we can also maximize the Chernoff distance [10] between the covariance matrices under both hypotheses at the output of the WT block. Since, in both hypotheses, the process has zero mean value, then the Chernoff distance is given by

$$\max_\xi \mu(\xi) = \max_\xi \left(\frac{1}{2} \ln \frac{|\xi C_{H_0} + (1-\xi)C_{H_1}|}{|C_{H_0}|^\xi |C_{H_1}|^{(1-\xi)}} \right), \; 0 \le \xi \le 1.$$

In the Gaussian case, the Chernoff distance leads to a closed-form expression of the Chernoff bound, that consists in an upper bound for the PE. Then, the value of the parameter a that leads to a maximum value of the Chernoff distance is expected to give a good separability between the two hypotheses, as it minimizes the Chernoff bound of the PE.

The likelihood ratio is established assuming that, under both hypotheses, the received process is Gaussian. Let us denote by X the vector obtained by prefiltering and decomposing the observation process under a set of orthonormal functions. The decision about which hypothesis the received signal belongs results from the Bayes test, that consists in the comparison of the *a-posteriori* probabilities:

$$P(\mathrm{H}_1|X) \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} P(\mathrm{H}_0|X). \quad (8)$$

For Gaussian signal and noise processes, the binary test (8) becomes

$$\mathcal{L} = X^\top (C_{H_0}^{-1} - C_{H_1}^{-1}) X \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} \ln\left(\frac{|C_{H_1}|}{|C_{H_0}|}\right) + 2\ln\left(\frac{P(\mathrm{H}_0)}{P(\mathrm{H}_1)}\right), \quad (9)$$

where $P(\mathrm{H}_i)$ is the, assumed known, *a priori* probability of hypothesis i, and $|\cdot|$ stands for the matrix determinant.

All the previous considerations were developed assuming that the nonlinearity parameter a is large, leading to near-Gaussian pdf in both hypotheses. However, simulation results show that the best processor (in the sense that it minimizes the PE) based on the proposed scheme, corresponds to a value of a that is close to the maximum standard deviation of the signal process, where strictly Gaussianity can not be assumed (although the signal information is not strongly degraded). This situation results from a compromise between two contradictory issues: in one hand, the process Gaussianity increases with increasing a, giving rise to a receiver close to the optimal; in the other hand, the hypotheses separability strongly decreases for large values of a.

3. SIMULATION RESULTS

For simulation purposes, we generate bandpass transients $s(i) = [p(i) \times \nu(i)] * f(i)$ ($*$ denotes the convolution) by gating and filtering independent Gaussian distributed sequences $\nu(i)$ with zero mean and spectral height 49, where

$$p(i) = 16.4 e^{-50(\frac{i}{207}-0.5)^2/2}, \quad \text{for } i \in [0; 207]$$

is the gating and $f(i)$ corresponds to the discretization, at the sampling rate of $20/207s$, of the continuous function filter

$$g(t) = IFT\left[\frac{rect(148, 223)(j\omega)^2}{(j\omega + 45 - 180)^2 (j\omega + 45 + 180)^2}\right].$$

IFT denotes the inverse Fourier transform, and $rect(x, y)$ is the rectangular function that takes the value one in the interval $[x, y]$ and zero elsewhere. The impulsive noise is generated as independent S-α-S noise with characteristic exponent $\alpha = 1.5$, location parameter $\ell = 0$ and dispersion $\gamma = 1$.

The WT decomposition is performed using Daubechies D14 filters [8], using only the coefficients at scale 2, where most of the energy of the transient signals lie.

Fig. 3 shows samples of both the generated signal, and the signal plus noise, and highlights the impulsive and spiky behavior of the S-α-S distribution. Fig. 4 presents the curves of the Chernoff distances computed from the approximated expression of eq. (7), and obtained from the correct computation of the covariance matrices under both hypotheses. It is interesting to point out that, when the parameter a takes values larger than the maximum standard deviation of the signal sequence $s(i)$ (corresponding to the case where assuming that $z(i)$ is independent of $s(i)$ is close to reality), then the approximation curve tends to the true one. In Fig. 5, we plot the receiver operating characteristics for 3 different values of a. The best case is close to the maximum obtained in fig. 4, thus showing, as expected, that the maximization of the Chernoff bound gives a good approximation to the minimization of the probability of error. Although the ROC shows a better performance for $a = 3$ when compared to the case where $a = 9$, the Chernoff distance in the latter case is larger. This contradictory result is due to the fact that, for $a = 3$, the observation process is far from being Gaussian distributed, and the Chernoff distance no longer reflects the class separability.

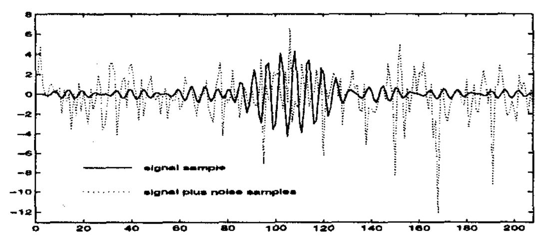

Figure 3. Signal and signal plus noise samples.

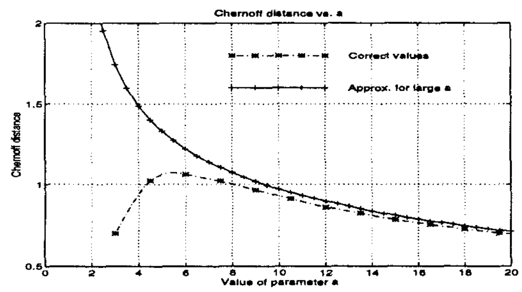

Figure 4. - Chernoff distances.

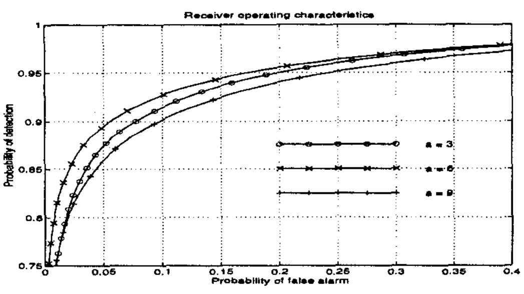

Figure 5. a) - Chernoff distances. b) - Receiver operating characteristics

4. CONCLUSIONS

The paper presents a detector of Gaussian bandpass transients under impulsive noise. The proposed scheme first performs a nonlinear memoryless filtering of the observation process, in order to obtain a finite-variance signal at its output. A nonlinearity parameter is tuned, either by looking at the receiver operating characteristic, or by an approximate method using the Chernoff distance, which requires a smaller computational load. Then, a linear decomposition is used, increasing the Gaussianity of the process under both hypotheses. For this purpose, the wavelet transform was chosen, due to its ability to reduce the complexity of bandpass processes, as well as its on-line algorithmic capabilities. At last, assuming Gaussianity, a quadratic likelihood ratio is computed, and compared with a threshold. Monte Carlo simulation results are presented using synthetic data.

REFERENCES

[1] M. Bouvet and S. C. Schwartz, "Comparison of adaptive and robust receivers for signal detection in ambient underwater noise," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 37, no. 5, pp. 621–626, May 1987.

[2] E. Voudouri and L. Kurz, "A class of sequential adaptive detectors for underwater non-gaussian noise environments," *IEEE Journal of Oceanic Engineering*, vol. OE-12, no. 1, pp. 38–47, January 1987.

[3] C. L. Nikias and M. Shao, *Signal Processing with Alpha-Stable Distributions and Applications*, Wiley interscience, 1995.

[4] P. Tsakalides and C. L. Nikias, "Maximum likelihood localization of sources in noise modeled as a stable process," *IEEE Transactions on Signal Processing*, vol. 43, no. 11, pp. 2700–2713, November 1995.

[5] F. M. Garcia and I. M. G. Lourtie, "A wavelet transform frequency classifier for stochastic transient signals," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, Atlanta, USA, May 1996, pp. 3058–3061.

[6] F. M. Garcia and I. M. G. Lourtie, "An arrival time estimator for transient stochastic signals using the wavelet transform," in *38th Midwest Symposium on Circuits and Systems*, Rio de Janeiro, Brazil, August 1995.

[7] S. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, July 1989.

[8] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Communications on Pure and Applied Mathematics*, vol. XLI, pp. 909–996, 1988.

[9] V. S. Pugachev, *Theory of Random Functions*, Pergamon Press, 1965.

[10] K. Fukunaga, *Introduction to Statistical Pattern Recognition, Second Edition*, Academic Press, 1990.