Robust Detection of Signal Classes

O.P. Kenny and L.B. White
1. Cooperative Research Centre for Robust & Adaptive Systems
2. Signal Analysis Discipline, Communication Division
Defence Science and Technology Organisation
PO Box 1500 Salisbury SA 5108
AUSTRALIA

Abstract

This paper proposes a robust detection statistic for signals whose parameters are uncertain. Standard detection schemes generally use time domain correlation which can be related to correlation based on the Wigner-Ville distribution by Moyal's identity. This paper shows that a more robust detection statistic is achieved by using generalised patterns in time-frequency space and deriving a non-linear time domain correlation.

The performance of the robust detection statistic is evaluated with the aid of *receiver operating curves*, two robust detection examples are given.

1.0 Introduction.

The signal processing requirement concerned within this paper is the robust detection of signals. With traditional detection scheme all the information about the process is assumed to be known, ie. both the deterministic and random natures of the process. Usually, detection schemes optimise the detection probability or the signal to noise ratio output with the result being the well known matched filter.

It is well known that the performance of the matched filter degrades dramatically when the signal deviates from the proposed model.

In reality, one does not always know the exact information about the process and so the detection algorithm must be able to give reasonable results under conditions that cause the signal to deviate away from the ideal conditions. In essence the detection algorithm needs to be inherently insensitive to signal model perturbations.

1.1 Signal space localisation.

It is well known characterisation of a non-stationary signal is possible by using a distribution that localises one of the signal's parameters, such as its energy, in both time and frequency. Probably the most well known of these representations is the short time Fourier transform which as been used widely in the analysis of speech.

Recognition and detection of a signal have been used in the past by time-frequency representations (TFRs) due to their ability to illustrate important information about the signal. The notion of robust time-frequency filtering enters naturally into detection applications because it is possible to generalise the signal space features. It has been shown that formulation of correlation type receivers is obtainable from TFRs due to Moyal's identity- a relationship between time-frequency correlations and time domain correlations.

A major problem associated with time-frequency correlation is that the dimensionality is increased by a factor of two in moving to the time-frequency space from the time domain. The two dimensional correlation that forms the basis of classification using TFDs are inherently computationally expensive, and limits real time capabilities.

This paper shows how the reduction of the computation load for time-frequency based correlation can be achieved by making use of eigenvalue decomposition.

2.0 Problem definition.

Consider a measurable signal which has finite energy with its inner product given by,

$$<x,y> = \int_{-\infty}^{+\infty} x(t)\, y^*(t)\, dt. \qquad (1)$$

and for any signals which are elements of $L_2(\mathbb{R})$ Schwartz's inequality holds

$$|<x,y>| \le \|x\|\, \|y\|. \qquad (2)$$

Both (1 & 2) form the basis of time domain matched filtering, and can be used to classify a measured signal from one of the specified nominal functions $y_1 \cdots y_N$. The decision rule in such a traditional classification procedure is to choose the signal which maximises the output of (1).

It can be shown that the inner product of the Wigner-Ville distribution (WVD) of the measured signal y(t) with the WVD of a known signal x(t) is related to the inner product defined in (1) [1], and so

$$\int_{R^2} W_x(t,f)\, W_y(t,f) dt\, df = |<x,y>|^2 \qquad (3)$$

where by definition, the WVD is

$$W_x(t,f) = \int_{-\infty}^{+\infty} x(t+\frac{\tau}{2})\, x^*(t-\frac{\tau}{2})\, e^{-i2\pi ft}\, dt. \qquad (4)$$

Using Moyal's identity the WVD can form the basis of a correlation classifier. To find the correlation of a signal based on the WVD it is necessary that the WVD of the signal under analysis be calculated.

Since by Moyal's identity, the two formulations for the correlation statistic are equal then one might asked the question- "why use time-frequency distributions at all to formulate the correlation statistic?" Replacement of the WVD by a more general form of time-frequency pattern function provides the answer to this question. This allows much greater flexibility in the classification process, as representation of a broader class is possible. The correlation statistic can now be redefined as,

$$\eta_j(x) = <W_x, \psi_j>, \quad j=1,\ldots,N. \qquad (5)$$

The computational complexity for such a correlation statistic is of order $O(T^2 \log T) + O(T^2)$, where T denotes the length of the signal. This is in contrast to the computation complexity for the one dimensional case that has order O(T).

As seen from these orders of complexity it would be advantageous if there were a reduction of the time-frequency processing requirements to one dimensional format, or a series

0-7803-2431-5/95 $4.00

of one dimensional operations so real time processing could be realisable.

3. 0 Theory of filter design based on time-frequency pattern function.

Generally, the windowing function $\psi_j(t,f)$ has Hermitian symmetry so that the bilinear time-frequency representation is real. The derivation that follows shows one can obtain a significant reduction in the computational complexity using eigenvalue decomposition of the time-frequency pattern function kernel. Truncation of this expansion allows the reduction in computational complexity since typical pattern classification functions will be of low effective rank. The kernel function is defined as

$$K_{Q_j}(t_1,t_2) = \int_{-\infty}^{+\infty} \psi_j(\frac{t_1+t_2}{2}, f)\, e^{i\pi f(t_1-t_2)}\, df. \qquad (6)$$

Rearranging and then substituting (4) into (5) yields,

$$\eta_j(x) = \int_{R^3} x(t+\frac{\tau}{2})x^*(t-\frac{\tau}{2})\, e^{-2\pi i f t}\, dt\, \psi_j^*(t,f)\, dt\, df \qquad (7a)$$

$$= \int_{R^2} x(t_1)x^*(t_2)\, K^*_{Qj}(t_1,t_2)\, dt_1 dt_2 \qquad (7b)$$

$$= \langle x, Q_j x \rangle. \qquad (7c)$$

where Qj denotes a mapping from $L_2(\mathbb{R}) \longrightarrow L_2(\mathbb{R})$ and is defined by,

$$(Q_j x) = \int_R K_{Q_j}(t,\tau)\, x(\tau)\, d\tau. \qquad (8)$$

The correlation statistic then has the expansion of the form

$$\eta_j(x) = \sum_{n=0}^{\infty} \lambda_n^{(j)} |\langle x, \zeta_n^{(j)} \rangle|^2. \qquad (9)$$

Where $|\lambda_0^{(j)}| \geq |\lambda_1^{(j)}| \geq \ldots$ are the eigenvalues of Qj, and $\zeta_n^{(j)} \in L_2(\mathbb{R})$ are the associated orthonormal eigenfunctions. Equation (9) follows since K_{Q_j} is Hermitian and admits the expansion [2]-[4]

$$K_{Qj}(t_1,t_2) = \sum_{n=0}^{\infty} \lambda_n^{(j)} \zeta_n^{(j)}(t_1)\zeta_n^{(j)*}(t_2) \qquad (10)$$

converging in $L_2(\mathbb{R}^2)$. The compact Fredholm operator generated by the kernel K_{Q_j} is defined by

$$(Q_j x)(t) = \sum_{n=0}^{\infty} \lambda_n^{(j)} \langle x, \zeta_n^{(j)}(t) \rangle \zeta_n^{(j)}(t). \qquad (11)$$

The correlation statistic now has the form defined by (12)

$$(Q_j x)(t) = \sum_{n=0}^{\infty} \lambda_n^{(j)} \langle x, \zeta_n^{(j)}(t) \rangle \zeta_n^{(j)}(t). \qquad (12)$$

For many practical applications, the pattern eigenvalues $\lambda_n^{(j)}$ will decrease rapidly in magnitude as $n \longrightarrow \infty$. Reduction of the computation required to obtain the correlation statistic is possible by taking into consideration only those eigenvalues that are dominant. The correlation statistic for the truncated set of eigenvalues is defined by

$$\eta_j^{(L)}(x) = \sum_{n=0}^{L-1} \lambda_n^{(j)} |\langle x, \zeta_n^{(j)} \rangle|^2. \qquad (13)$$

The error in using the reduced order filter instead of the true correlation statistic is given by

$$|\eta_j - \eta_j^{(L)}| \leq \|x\|^2 \, [\sum_{n=L}^{\infty} |\lambda_n^{(j)}|^2]^{1/2}. \qquad (14)$$

This bound can be obtained by subtracting the kernel function defined by (10) from the true kernel function expansion.

$$K_{Qj}^{(L)}(t_1, t_2) = \sum_{n=0}^{L-1} \lambda_n^{(j)} \zeta_n^{(j)}(t_1) \zeta_n^{(j)*}(t_2) \qquad (15)$$

where L denotes the lower rank approximation to K_{Q_j} then

$$\|K_{Qj} - K_{Qj}^{(L)}\|^2 = \sum_{n=L}^{\infty} |\lambda_n^{(j)}|^2. \qquad (16)$$

Since the error between the true and reduced correlation statistic is given as

$$|\eta_j - \eta_j^{(L)}| = |\langle x, Q_j x \rangle - \langle x, Q_j^{(L)} x \rangle| \qquad (17a)$$

$$\leq \|x\|^2 \|Q_j - Q_j^{(L)}\| \qquad (17b)$$

Note from (17b) that the error bound is only dependent on the norm of x. Consequently, the time-frequency correlator can be approximated by a bank of one dimensional correlators.

Although the eigenfunction sequences must be stored to obtain the correlation statistic and for large data lengths this may become cumbersome. State space realisation of the filters can be achieved which reduce the storage requirements.

4.0 Derivation of analytic receiver operating curves.

To obtain some performance indicators for the proposed correlation statistic, analytic performance equations are derived. The probability of detection and false alarm completely characterises the performance of the robust detection algorithm, and are used to form analytic receive operating curves.

Before deriving the required probability density function for the non-linear detection statistic some preliminary assumptions are made about the processes involved. It is assumed that all possible waveforms $x_i(t)$ can be decomposed into a set of orthogonal functions and projection of the noise component onto this complex vector space gives in phase and quadrature components.

The probability density function of the projected noise component has a Rayleigh probability density function for its magnitude and a uniform probability density function for the phase.

Consider now the robust detection statistic where the signal under detection that has been corrupted by noise. The output of the detector is given by,

$$\eta = \sum_{i=1}^{N} \lambda_i \left|\langle (x+n), \phi_i \rangle\right|^2 \qquad (18a)$$

$$= \sum_{i=1}^{N} \lambda_i \left\{ \left(x_i^c + n_i^c\right)^2 + \left(x_i^s + n_i^s\right)^2 \right\}. \qquad (18b)$$

The aim is to determine the probability density function of the detection statistic and the probability of false detection which requires in both cases the evaluation of the probability density function $p_n(\alpha;\ \delta;\ y)$ that has a linear quadratic form

$$y = \sum_{i=1}^{2N} \alpha_i \left(Z_i + \delta_i\right)^2 \qquad (19)$$

where Z_i are mutually independent standardised normal variables. The analytic evaluation of such a linear quadratic form has been the subject of many papers. The analytic expression for the probability of a linear quadratic form is not trivial. The expression given here can be found in the paper by Kotz et al [5]. The first step is to obtain the Laplace transform of the overall probability density function of the linear quadratic form (LQF), given by (20).

$$L_T(s) = \exp\left\{-\sum_{i=1}^{N} \frac{\delta_i^2 \alpha_i s}{\left(1 + 2\alpha_i \sigma_i^2 s\right)}\right\} \prod_{j=1}^{N} \left(1 + 2s\alpha_i \sigma_i^2\right)^{-\frac{1}{2}}. \qquad (20)$$

What is required is an expansion of the probability density function of the form,

$$p_n(\alpha;\ \delta;\ y) = \sum_{k=0}^{\infty} a_k^c\, \beta^{-1} \chi\left(n+2k;\ y/\beta\right). \qquad (21)$$

where

$$P_n(\alpha;\ \delta;\ y) = \sum_{k=0}^{\infty} a_k\, \Theta^k \qquad (22)$$

if $\Theta = \Gamma(s)$ this implies that $s = \varphi(\Theta)$ which enables a moment generating function to be defined as,

$$M(\Theta) = L_T\left(\alpha;\ \delta;\ \varphi(\Theta)\right) / \xi\left[\varphi(\Theta)\right] = \sum_{k=0}^{\infty} a_k\, \Theta^k. \qquad (23)$$

By choosing a χ^2 distribution expansion,

$$\xi(s) = \left(1 + 2s\beta\right)^{-n/2} \quad \Gamma(s) = \left(1 + 2s\beta\right)^{-1}$$

$$\varphi(\Theta) = \left(1 - \Theta\right)/2\beta\Theta,\ \xi\left[\varphi(\Theta)\right] = \Theta^{n/2},\ \gamma_j = 1 - \beta/\alpha_j \qquad (24)$$

and then substituting into the moment generating function gives

$$a_0^c = M(0)\, A\, e^{-v/2} \qquad (25)$$

with $v = \sum_{i=1}^{n} \delta_i^2$,

$$a_k^c = k^{-1} \sum_{r=0}^{k-1} O_{k-r}^c\, a_r^c \qquad (26)$$

$$O_k^c = \frac{1}{2} k \sum_{j=1}^{n} \delta_j^2 \gamma_j^{k-1} + \frac{1}{2} \sum_{j=1}^{n} \left(1 - k\delta_j^2\right) \gamma_j^k \quad k \geq 1. \qquad (27)$$

The coefficient used for the analytic receiver operating curves where generated by a program written in matlab. The examples which follow in the next section show the advantages and the cost incurred in using a robust detection filter.

5.0 Application of robust detection.

In this section two examples are given to demonstrate the use of robust detection of signals. The robust detection scheme is compared to that of a matched filter. So that a fair comparison is made between these detection approaches the outputs of the matched filter and the robust detector are normalised so that each has the same noise power output.

5.1 Robust detection of a sinusoidal signal.

In this example the robust detection scheme is used on a sinusoidal function whose value of frequency is not exactly known, but assumed to lie within a frequency region.

Figure 1 shows a comparison between the robust detector and the matched filter output as a function of the deviation from the nominal frequency. It can be seen that when the nominal frequency on which the matched filter was designed varies then the performance of the matched filter degrades. The output of the robust detection for the same noise power output does not perform as well as the matched filter when the frequency of the signal is near nominal frequency, but still performs reasonably well when the frequency deviates from its nominal frequency.

To see the effect of varying the frequency of the sinusoidal signal the corresponding AROC where obtained. The deviation frequency values where 0 (corresponding the nominal frequency), .25, .5, .75, 1. The results of the experiment where plotted in figure 2. As shown in this figure the effect of varying the frequency away from the nominal frequency degrades the performance for the matched filter. This can be seen as the probability of detection deceased and the probability of false detection increased with the AROC approaching the diagonal line (equal probability of false detection and detection) as the deviation frequency increased .

In comparison the AROC were plotted on the same graph for the robust detector, shown as a '+' symbol. The deviation frequencies where the same, however the results show that the robust detector was more tolerant to variations to the signal frequency.

5.2 Robust detection of a linear FM signal with unknown chirp rate.

In this example the signal under detection is a linear FM signal. Here the exact value of the chirp rate is not known, but assumed to lie within an interval, where α_0 and α_1 denote the upper and lower chirp rates respectively. A matched filter was designed for the case where the value of the chirp rate was α_0. Figure 3 shows the outputs for both the matched filter and the robust detector. Here as the signal's chirp rate increases it can be seen that the performance of the matched filter degrades, whereas the robust output still gave reasonable results.

The nominal chirp rate in the experiment was .25 Hz/sec and the chirp rate was increased in steps of .005 until the chirp rate was .27 Hz/sec. The AROC, figure 4, for the robust detector plotted on the same graph shows that for all the variations in the chirp rate of the signal the performance of the robust detector stayed the same.

6.0 Conclusions.

This paper has proposed a robust detection statistic for signals whose parameters are uncertain. It was shown that a more robust detection scheme can be achieved by a generalised time-frequency pattern, that defines a class of detectable signals, was decomposed into a set of orthonormal eigenfunction and so by considering only the dominant eigenvalues a reduced order correlation statistic was obtained. This resulting statistic can not be otherwise implemented by a corresponding linear process illustrating the important feature that components inherent in the robust detector interact with other to produce the desired response.

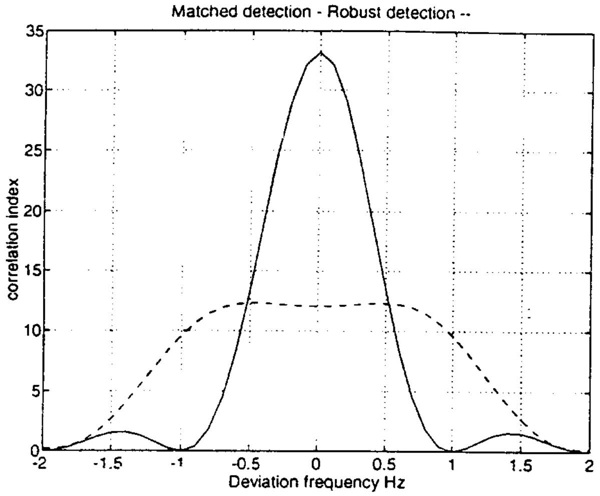

Figure 1 Detector outputs for Example 1.

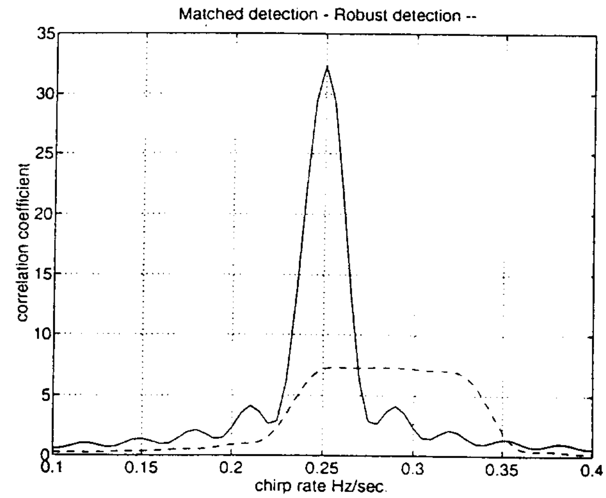

Figure 3 Detector outputs for Example 2.

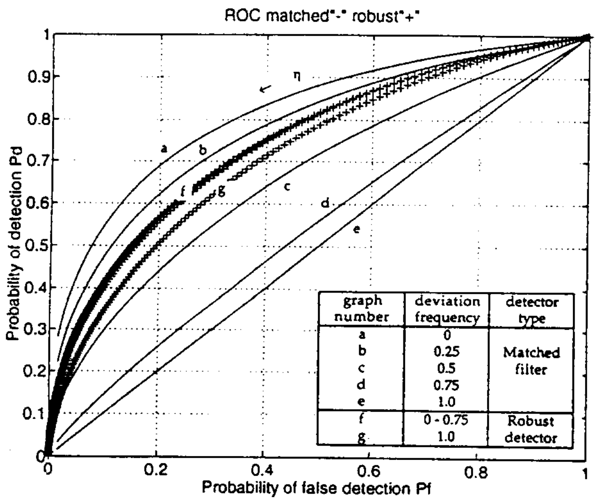

Figure 2 AROC Curve for Example 1.

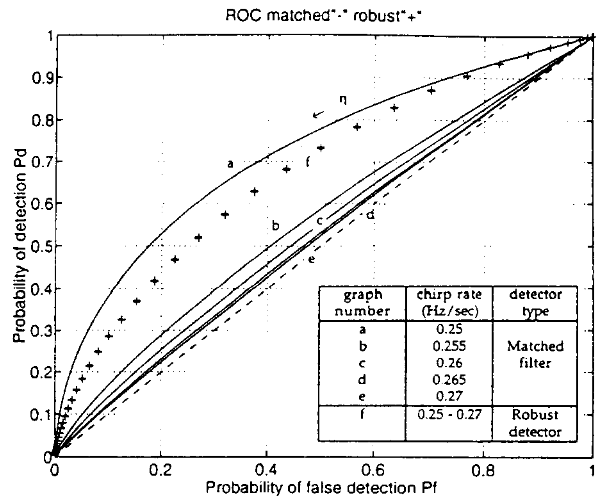

Figure 4 AROC Curves for Example 2.

The performance of the robust detection statistic was also evaluated with the aid of *receiver operating curves* which evolved the analytic derivation of the probabilities of detection and false alarm based on central chi squared expansion of a linear quadratic form.

It was shown that the robust detection statistic under nominal conditions did not perform as well as standard detection scheme, as one would expect, but gave better performance when the parameters of the signals deviated away from its nominal conditions. More detailed information about this work can be found in [6].

Acknowledgements: The authors thank the Australian Cooperative Research Centre for Robust and Adaptive Systems for their financial support.

7.0 References.

[1] L. Cohen, "Generalized phase space distribution functions", *J. Math. Phys.*, Vol. 7, pp. 781-786, 1966.

[2] N.M. Marinovic and G. Eichmann, "An expansion of Wigner distribution and its applications", *IEEE Int. Conf. Acoustics, Speech and Signal Processing,* 1985, pp. 1021-1024.

[3] L.B. White, "Transition kernels for bilinear time-frequency distributions" *IEEE trans. Signal Processing,* v. 39, no. 2, 1991.

[4] M.G. Amin and M.T. Schiavoni, "Time-varying spectrum estimation via multidimensional filter representation", *SPIE Advanced Algorithms and Architectures for Signal Processing,* Vol. 1152, San Diego, 1989.

[5] S. Kotz, N.L. Johnson and D.W. Boyd, "Series representation of distributions of quadratic forms in normal variables II. non-central case", *Ann. of statistical Math.*,pp 838-848, 1966.

[6] O.P. Kenny and L.B. White, "Roust detcetion of signal classes", submitted to *IEEE Trans on Signal Processing,* 1994.