A Methodology for Detection and Classification of Some Underwater Acoustic Signals Using Time-Frequency Analysis Techniques

BOUALEM BOASHASH, SENIOR MEMBER, IEEE, AND PETER O'SHEA, STUDENT MEMBER, IEEE

***Abstract*—Signal detection techniques based on time-frequency signal analysis with the Wigner–Ville distribution (WVD) and the cross Wigner–Ville distribution (XWVD) are presented. These techniques are shown to provide high resolution signal characterization in a time-frequency space, and good noise rejection performance. This type of detection is applied to the signaturing, detection, and classification of specific machine sounds: the individual cylinder firings of a marine engine. For this task, a four step procedure has been devised. 1) The autocorrelation function (ACF) is first employed for ascertaining the number of engine cylinders and the firing rate of the engine. 2) Further correlation techniques are then used to detect the time at which individual cylinder firing events occur. 3) WVD and XWVD based analyses follow to produce high resolution time-frequency signatures. 4) Finally, 2D correlations are employed for classification of the individual cylinders. The proposed methodology is tested on real data. XWVD based detection is also applied to detection of a transient with unknown waveshape (using real data).**

I. Introduction

THE detection of transients in noise is a problem of major importance in underwater acoustics, seismic processing, and radar. In these applications, the transient is to be detected and its beginning and end determined. In some circumstances, when a set of reference templates is available, a second stage of processing is needed to perform classification.

A particular example where this type of detection is important is an underwater surveillance system, where the detection of transients can signal the presence of an underwater vessel. In many situations, no reference templates are available for comparison; this problem of detecting transients of unknown waveshape is considered in Sections II-A, III, and V of this paper. Ideally, however, one would have a library of profiles for all possible vessels, with each profile containing as much information as possible for the task of discrimination. One such piece of information is the power spectral density (PSD), which gives indications of the engine's structure. Others could be characteristic "signatures" of particular transient activity unique to that vessel. In this paper, signaturing, using time-frequency analysis with the Wigner–Ville distribution (WVD) and cross Wigner–Ville distribution (XWVD) is considered, with specific application to classification of individual cylinder firings. In this scheme, the ACF is used as an aid for determining the engine firing rate and number of cylinders. The WVD is then used successfully to form time-frequency signatures of the individual cylinders of a marine engine, with these being used for subsequent detection and classification. An example of application of XWVD based techniques to detection of an unknown transient is also provided. The transient was from an unknown source in a marine diesel. The data was provided by the Australian Defence Science and Technology Organisation (DSTO).

Manuscript received December 10, 1988; revised November 7, 1989. This work was supported by the Australian Defence Science and Technology Organisation under Research Agreement CDS 10/87-DST 86/726.

B. Boashash is with CRISSP, University of Queensland Brisbane, Australia 4072, and with the Graduate School of Science and Technology, Bond University, Gold Coast, Australia, 4229.

P. O'Shea is with CRISSP, University of Queensland, Brisbane, Australia 4072.

IEEE Log Number 9038418.

II. Detection and Classification of Transients

A. Detection

To solve the problem of detecting a signal $s(t)$, having unknown waveshape, a number of approaches have appeared in the literature. The problem is to determine, in a noisy environment, whether a signal is present (hypothesis H_1) or not (hypothesis H_0)

$$H_0\colon\ r(t) = n(t)$$
$$H_1\colon\ r(t) = s(t) + n(t) \tag{1}$$

where $n(t)$ is an additive noise process. Some of the earliest detectors were the energy detector [1] and the spectral density correlator [2]. More recently, adaptive detectors have been proposed. For these, a minimum variance estimate of the signal is first obtained, then this estimate is used as the reference signal for a conventional detector [1]. One method for performing this critically important estimation in a background of white Gaussian noise involves modeling the signal as the impulse response of a filter having a rational system function [3]. It was shown in [3] that this adaptive detector is better than the energy detector, at least when the noise variance is known and when the number of parameters to be estimated is small compared with the number of data points ("better" in the sense of having higher detection probability for a given false alarm probability).

B. Classification

Where classification is required as well as detection, one must first segment the signal by determining the tran-

0096-3518/90/1100-1829$01.00

sient's beginning and end. AR modeling is often used [4], [5]. Because of the short duration of transients, parametric modeling has also often been used to represent the information content of the transient. In such cases, the feature vectors can contain the model parameters, along with other variables such as the transient duration, the number of subevents, instantaneous amplitude, instantaneous phase, and various moments of the envelope and frequency histogram data [4], [6]. High order AR parameters were used as features for a classification procedure in [6], which was based on a minimization of the "spectral distance" between the unknown transient and any one of a set of known ones.

III. Signal Detection Using the WVD

While the model based approaches presented in [3] and [4] perform well under a number of circumstances, their performance is limited for nonstationary signals in high noise environments, particularly when the arrival times of the transients are unknown. For this, and other reasons, several authors have proposed detection and estimation methods based on time-frequency distributions. Spectrogram correlators, for example, have been in use for transient analysis for some time and are reviewed in [7]. While they provide very good intuitive representations of the signal's spectral content, resolution is poor, the detector is suboptimal due to the loss of phase information, and reconstruction of the signal often occurs at some cost.

The Gabor representation has also been used to detect transients of unknown waveshape and arrival time [8]. This representation was chosen largely because of its inherent ability to localize signals in time and frequency. The authors of [8] chose a one sided Gaussian window to match the often abrupt changes characteristic of transients and formed a detection statistic which is a function of the Gabor coefficients.

The Wigner–Ville distribution, a member of Cohen's class of generalized time-frequency representations [9], has been proposed by several authors as a tool for the detection problem [10]–[12] and [15]. It is defined by

$$W_{z_s}(t, f) = \int_{-\infty}^{+\infty} z_s(t + \tau/2) z_s^*(t - \tau/2) e^{-j2\pi f\tau} \, d\tau \tag{2}$$

where $z_s(t)$ is the analytic signal corresponding to $s(t)$ [10].

Similarly, the XWVD of two signals, $s(t)$ and $r(t)$, is

$$W_{z_r}(t, f) = \int_{-\infty}^{+\infty} z_s(t + \tau/2) z_r^*(t - \tau/2) e^{-j2\pi f\tau} \, d\tau \tag{3}$$

where $z_r(t)$ is the analytic signal corresponding to $r(t)$.

One of the reasons for using the WVD (XWVD) for detection purposes stems from the equivalence between time domain correlations and XWVD correlations. The time domain correlations, which appear as the solutions to many classical detection problems, may thus be replaced by equivalent WVD (XWVD) correlations [16]. The advantage of this replacement is that time-frequency feature isolation and consequent noise suppression using time-varying filtering is easier [10], especially where the signal waveshape is unknown. Classical detectors and estimators have used autocorrelation functions to represent nonstationary processes. The Wigner–Ville spectrum (which is simply the expected value of the Wigner–Ville distribution) of a nonstationary process has been shown to contain essentially the same information as its autocorrelation function (they form a Fourier transform pair) [10], but describes time-varying signals well. We are led naturally, then, to substitute the WVD for the autocorrelation function, and the XWVD for the cross-correlation function.

A. Detection with the WVD and the XWVD

Theoretical results for detection of a signal using the Wigner distribution (WD), as well as for the cross Wigner distribution (XWD) were derived by Kumar and Carroll in [12], [13], and compared with matched filter detection. Their results are summarized here and extended for application to the WVD, which uses the analytic signal rather than the real signal, and therefore provides fundamental and practical advantages over the WD [17].

Consider a signal $s(t)$ with energy $= A$, sent through a noisy transmission channel, so that the output of the channel is $r(t) = s(t) + n(t)$, where $n(t)$ is a zero mean, white Gaussian noise process of spectral height N_0. A detection statistic (η) is formed, such that if η exceeds a certain threshold, the signal is considered present; if less, it is decided that no signal is present. A signal-to-noise ratio (SNR), which provides a measure for comparison between WVD, XWVD, and the matched filter detection methods, is defined in [12] as

$$\mathrm{SNR} = \frac{\left| E(\eta | H_1) - E(\eta | H_0) \right|}{\left\{ 1/2 \left[\mathrm{var}\,(\eta | H_1) + \mathrm{var}\,(\eta | H_0) \right] \right\}^{1/2}} \tag{4}$$

where $E(\eta | H_1)$ and $E(\eta | H_0)$ are the expected values of η given H_1 and H_0, respectively, and var $(\eta | H_1)$ and var $(\eta | H_0)$ are the variances of η given H_1 and H_0, respectively.

1) Detection with a Matched Filter: For a matched filter we have

$$\eta_{\mathrm{mf}} = \int_{-\infty}^{+\infty} z_r(t) z_s^*(t) \, dt = \langle z_r z_s \rangle \tag{5}$$

where $\langle z_r z_s \rangle$ denotes the inner product of $z_r(t)$ and $z_s(t)$, as defined by (5), and

$$\mathrm{SNR} = \sqrt{A/N_0}. \tag{6}$$

2) Detection with the WVD: For WVD based detection

$$\eta_{\mathrm{wvd}} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} W_{z_r}(t, f) W_{z_s}^*(t, f) \, dt \, df. \tag{7}$$

Now Moyal's formula [18] states that for two XWVD's

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} W_{z_1z_2}(t,f)\,W^*_{z_3z_4}(t,f)\,dt\,df = \langle z_1z_3\rangle\langle z_2z_4\rangle^*. \quad (8)$$

Hence (7) reduces to

$$\eta_{\text{wvd}} = \left|\langle z_rz_s\rangle\right|^2 \quad (9)$$

and the noise performance for WVD detection is given by [12]

$$\text{SNR} = \sqrt{A/N_0}\cdot 1/\sqrt{1+N_0/A}. \quad (10)$$

This expression equals the SNR for a matched filter scaled by a factor of $1/\sqrt{1+N_0/A}$. For small values of A/N_0, the SNR is substantially reduced from the matched filter case. This reduction is due to the nonlinearity of the WVD which accentuates the effects of noise by producing both autoterms and cross-terms of the noise in its spectral formulation. That is, the WVD propagates noise from one region of the t-f plane to another, even showing noise at a particular time when there is none present [14]. However, even for low SNR, estimators based on the WVD can be more useful than time-domain based estimators, since they allow a "built-in" time-varying filtering operation not possible in the time domain (see Sections IV and V). Furthermore, the noise propagating properties of the WVD can be reduced by windowing the distribution.

3) Detection with the XWVD: The XWVD based detection statistic is

$$\eta_{\text{xwvd}} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} W_{z_rz_s}(t,f)\,W_{z_s}(t,f)\,dt\,df. \quad (11)$$

Using Moyal's formula [18], this statistic becomes

$$\eta_{\text{xwvd}} = \langle z_rz_s\rangle\langle z_sz_s\rangle^* = A\langle z_rz_s\rangle. \quad (12)$$

The XWVD detection statistic is thus the matched filter detection statistic multiplied by the signal energy (i.e., by a constant). Hence, the SNR for XWVD detection will be identical to that of the matched filter.

Hence, the detection statistics formed by the correlation of the reference signal WVD and the XWVD of the reference and observed signals is the same as that of the standard cross correlator except for a constant. The former statistic may then be used equivalently in any detection schemes. We will see later that due to the extra degree of freedom given by the 2D nature of the WVD, XWVD based estimators can perform better if this extra flexibility is used.

B. Interpretation and Application of XWVD Scheme

In general, the XWVD is complex, becoming the WVD (which is real) for $z_s = z_r$. Thus any distortion of the signal other than a real valued scaling causes real and imaginary oscillating "noise" components to appear in the XWVD. When a 2D correlation with the reference WVD is performed over the time-frequency plane, the imaginary part is rejected, and additionally, almost all the real oscillations are smoothed out, which accounts for the improved noise performance.

For practical purposes one may also form a "modified XWVD" detection statistic

$$\eta_m = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} W_{z_xz_s}(t,f)\,W^*_{z_xz_r}(t,f)\,dt\,df \quad (13)$$

$$= \langle z_xz_x\rangle\langle z_sz_r\rangle^* = A_x\langle z_rz_s\rangle \quad (14)$$

where A_x is the energy of the arbitrary signal $z_x(t)$.

This statistic is again equivalent to the cross-correlator statistic, so that it is possible to form a class of optimal 2D correlation statistics by varying $z_x(t)$. The usefulness of the class of detectors specified by (13) is that the estimation of $s(t)$ has now been replaced by estimation of $W_{z_xz_s}(t,f)$. Thus the estimation problem has been transferred into a time-frequency space, where feature selection and time-varying filtering are easier. Normally, in the WVD (XWVD) plane, the presence of cross terms makes it very difficult to distinguish multicomponent signals from the noise, prohibiting the appropriate filtering. If, however, $z_x(t)$ is chosen wisely, the autoterms and the cross-terms can be isolated, allowing relatively easy 2D feature extraction with built-in noise suppression.

C. Signal Estimation Examples

Previous work has demonstrated that for monocomponent asymptotic FM signals, the instantaneous frequency (IF) law $f_i(t)$ is a very useful characteristic of the signal for detection [10]. Since the signal contribution is concentrated around $f_i(t)$, the application of a 2D windowing which preserves all the points (t,f) in the neighbourhood $f_i(t)$ and filters out the others preserves the useful information in the signal and indeed increases by an order of magnitude the SNR. The subsequent application of a 2D cross correlation should then enhance the detection scheme [11].

According to the theory in the previous section, the modified XWVD scheme is used in preference to the WVD based one. The signal $z_x(t)$ referred to in (13) is taken to be a unit amplitude signal, reconstructed from the IF estimate of the signal

$$z_x(t) = \Pi_T(t - T/2)e^{j2\pi\Phi(t)} \quad (15)$$

where $\Phi(t) = \int_0^t f_i(\alpha)\,d\alpha$ and $\Pi_T(t) = 1$ for $-T/2 \le t \le T/2$ and 0 elsewhere.

From this signal, one can construct the cross Wigner-Ville distribution $W_{z_xz_r}(t,f)$. See, for example, Fig. 1 which shows the WVD of a noisy linear FM signal, and Fig. 2 which shows the XWVD (magnitude) of the same signal, with the reference being formed as in (15). Noise reduction through thresholding or windowing may then be affected to yield the estimate $\hat{W}_{z_xz_s}(t,f)$ in Fig. 3. This estimate can be used in (13) to form the appropriate detection statistic η_m. Note that what is particularly useful about this approach is the ability to localize the signal in

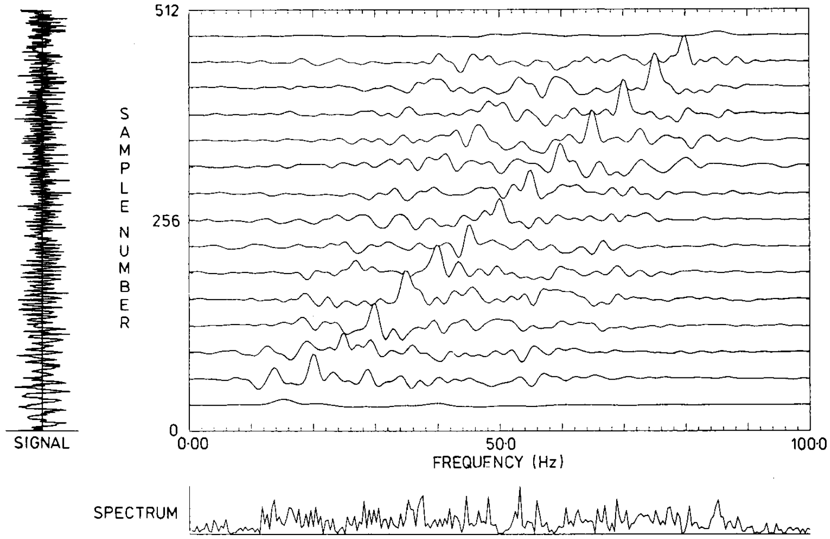


Fig. 1. WVD of linear FM signal in 0-dB noise.

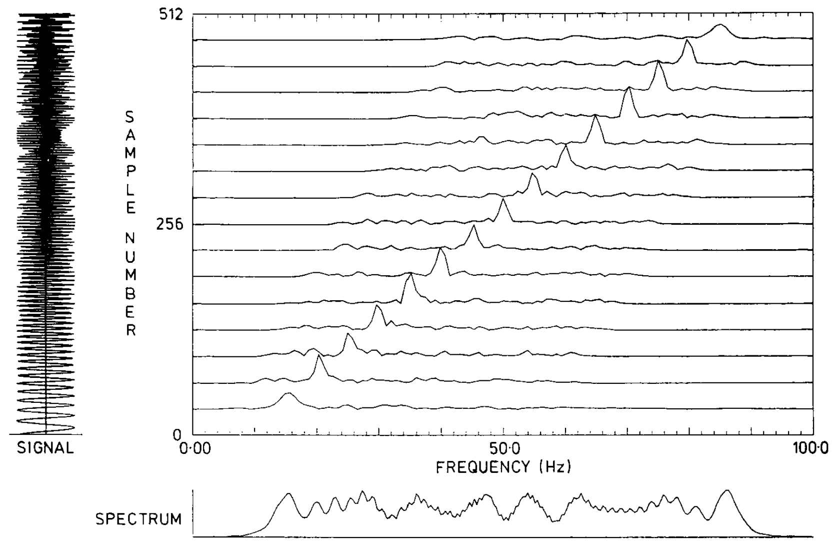


Fig. 2. XWVD of linear FM signal in 0-dB noise.

the time-frequency plane, so that the noise (which is dispersed broadly over the time-frequency plane) can be largely eliminated.

The method may be generalized to multicomponent signals, by forming a $z_x(t)$, such that

$$z_x(t) = \sum_{i=1}^{n} z_{x_i}(t) \tag{16}$$

where the $z_{x_i}(t)$ represent the instantaneous frequency laws of the various components. The procedure for estimating $\hat{W}_{z_x z_x}(t, f)$ then involves estimation of the individual $\hat{W}_{z_{x_i} z_x}(t, f)$ and summing

$$\hat{W}_{z_x z_x}(t, f) = \sum_{i=1}^{n} \hat{W}_{z_{x_i} z_x}(t, f). \tag{17}$$

With this approach, the cross-terms can be isolated from the autoterms, with resultant noise suppression and time-frequency feature selection being relatively easy. Fig. 4 shows the WVD of three linear FM signals of unknown slope in 3-dB noise. Fig. 5 shows a $\hat{W}_{z_x z_x}(t, f)$ estimate, where the frequency laws have been estimated by an algorithm based on dechirping techniques [19]. Note that

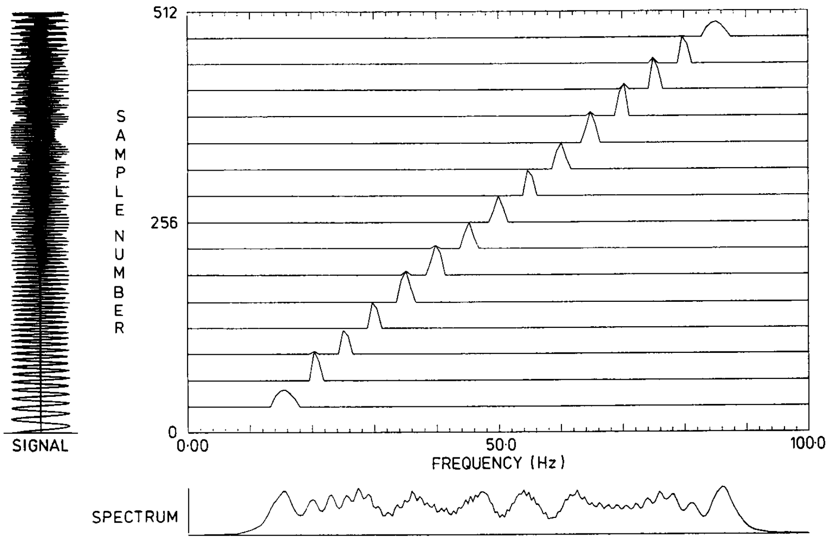


Fig. 3. Windowed XWVD of linear FM signal in 0-dB noise.

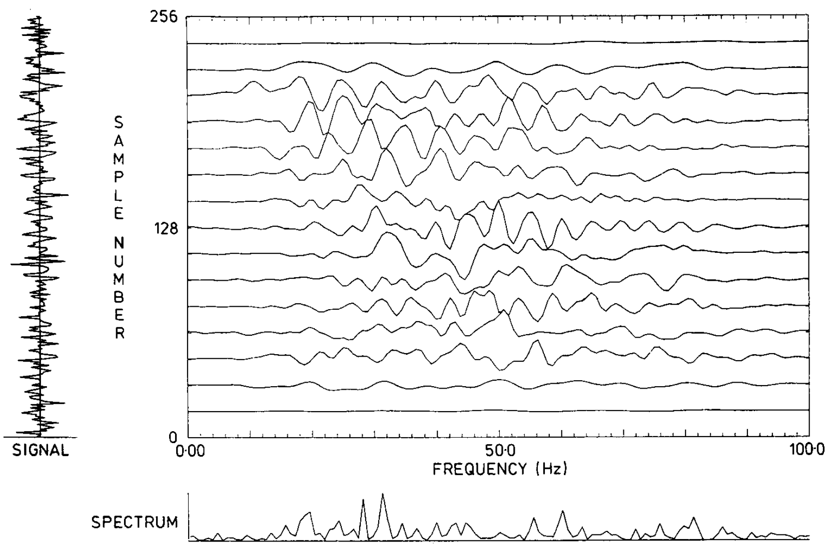


Fig. 4. WVD of 3 linear FM signals of unknown slope (3-dB noise).

here the cross-terms have been discarded to achieve good noise suppression. In some situations, one may not need to specifically estimate the laws at all, but may rely on data averaging and feature selection techniques. Such an example is presented in Section IV.

D. Digital Implementation

A degradation in SNR of up to 0.5 dB was reported in [13] when using the discrete WD for detection, as opposed to the continuous time WD. It was shown explicitly in [20] and implicitly in [21] that in using the WVD there is no SNR degradation. The degradation of the WD scheme as presented in [13] is due to the analytic signal not being used in the formulation of time-frequency representation, with aliasing occurring as consequence. As pointed out in [17], the practical use of the WD for time-frequency signal analysis must involve calculation of the analytic signal.

E. Signal Classification Using the WVD

The main advantage of classification techniques based on time-frequency distributions such as the WVD (XWVD), and as proposed in this paper, is the greater potential for formation of the feature vector, which is af-

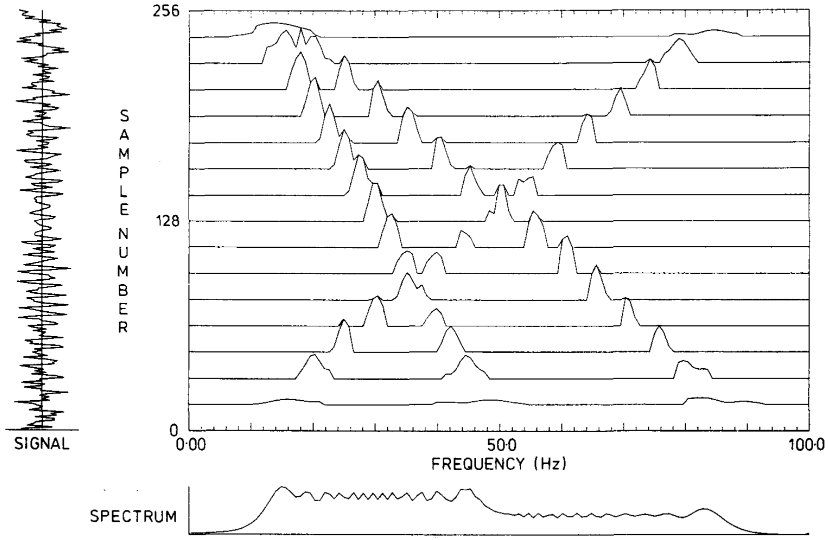


Fig. 5. XWVD estimate of 3 FM signals (artifacts eliminated).

forded by the 2D nature of the representations. If, for example, one knows that a certain transient has n time-frequency features, one can form a number of individual estimates $\hat{W}_{z_i z_i}(t, f)$, with the overall estimate being given by their sum according to (17).

IV. Application I—Signaturing, Detection, and Classification of Individual Cylinder Firings

In this section, we develop a methodology for detection and classification of particular underwater transients: the individual firings of an engine. The aim is to classify a given firing according to the cylinder which produced it. The application highlights the concepts outlined in previous sections. A four step procedure has been devised to this effect.

A. Step 1. Estimation of the Number of Cylinders and Engine Firing Rate

This first step is a preprocessing stage used for segmentation of the engine signal into individual cylinder firings. The segmentation techniques referred to in Section II-B are not appropriate here, since we are not considering isolated transients, but rather, "repeating" transients. Performing this segmentation requires knowing the engine's rate of firing as well as the number of cylinders. The method for predetermining them is based on the engine's autocorrelation function (ACF). The following terms are used throughout the text: CSR = the Crankshaft rotation rate (i.e., engine speed); CFR = the cylinder firing rate (firing rate of a particular cylinder); EFR = engine firing rate.

For a four stroke engine, we have CFR = CSR/2 and EFR = CFR × number of cylinders.

Several engines were used to provide test data to "calibrate" this technique for determining the EFR and number of cylinders. One example used was a 4 cylinder diesel tractor engine running at 1100 r/min. For this case, the engine speed (CSR) is given by CSR = 1100/60 = 18.33 Hz. Since the engine is four-stroke, the CFR is half this amount, while the EFR is four times the CFR, i.e.,

$$\text{CFR} = \text{CSR}/2 = 9.16 \quad \text{Hz} \quad \text{and}$$

$$\text{EFR} = 4 \cdot \text{CFR} = 36.66 \quad \text{Hz}.$$

The PSD for this sound is seen in Fig. 6 and its ACF in Fig. 7. The EFR from either the PSD or the ACF is found to be 36.3 Hz. Strong harmonics of the EFR are also evident in the PSD. The CFR spacing (indicated on the PSD and ACF) is found to be 9.2 Hz, and again seems to correlate with the calculated value. The number of cylinders is determined either by dividing the EFR by the CFR or, more easily, by inspection from the ACF (by noting the obvious periodicities). The ACF also shows every second peak to be stronger than the surrounding ones. This pattern is due to the periodicity of two cylinder firings (for a 4 cylinder 4-stroke) corresponding to the engine speed.

Note that while there is theoretically no difference in information content between the two representations, the PSD and ACF, apart from any windowing which may be applied in forming them, the autocorrelation can often show up periodicities in a way which is easier to interpret [22]. This can be clearly seen in Figs. 8 (autocorrelation for a length of signal from an underwater diesel sound) and Fig. 9 (PSD for the same signal segment). The periodicity every fourth cycle is much easier to detect in Fig. 8 than it is in Fig. 9. The autocorrelation function, then, can be a more natural tracker of short-term periodicities, and is thus preferred in this preprocessing stage.

The ACF of another machine sound is shown in Fig. 10. It corresponds to a 3 cylinder tractor engine. This fig-

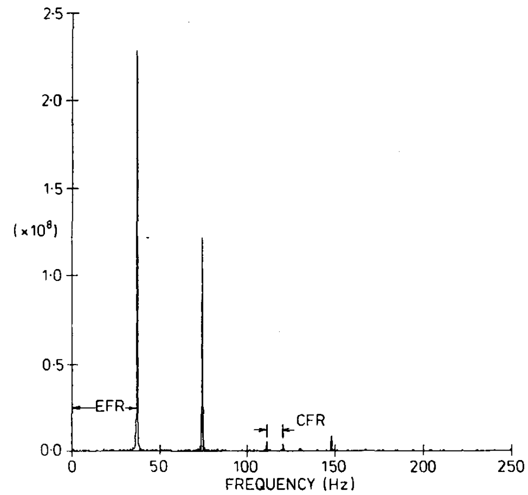


Fig. 6. PSD of 4 cylinder diesel (linear scale).

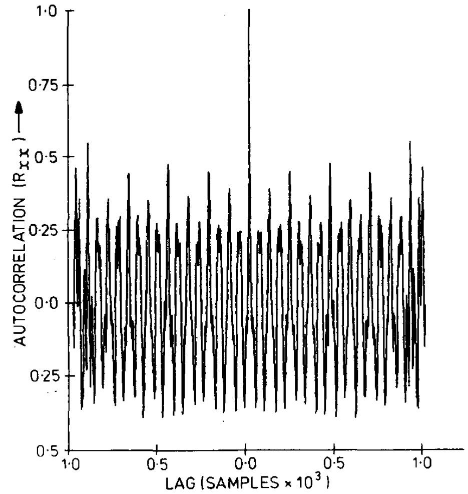


Fig. 8. Autocorrelation of underwater diesel.

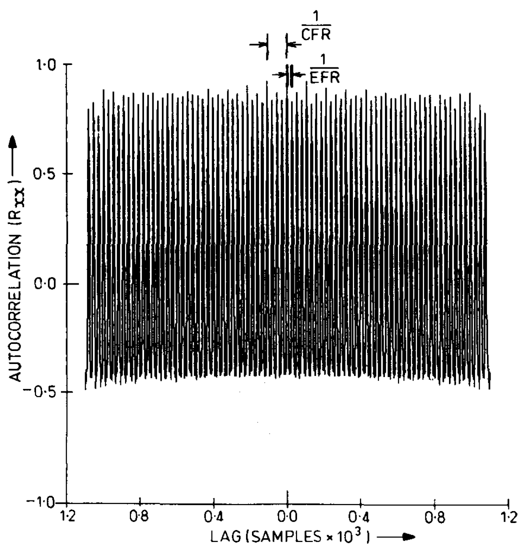


Fig. 7. Autocorrelation of 4 cylinder diesel.

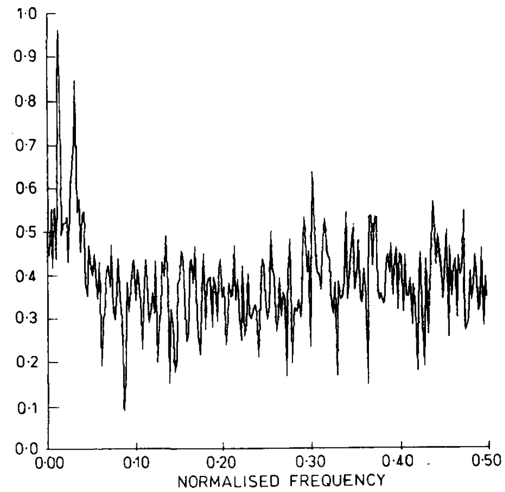


Fig. 9. Log power spectrum of underwater diesel.

ure shows a sequence of peaks corresponding to the cylinders firing, with very obvious major peaks appearing at every third one. A number of other engines were tested and, for simple engine sound signals, the EFR and the number of cylinders were easily obtained. Problems arose, however, when the engines became larger and more complicated.

B. Step 2. Detection

In this and the subsequent sections, we deal with the problem of determining whether a particular cylinder firing could be detected and identified (classified) from a given set of signatures. To determine whether a cylinder firing event has occurred and at what time, we examine the ACF once more. Examination of Figs. 7–9 shows that cylinder firing events are detected quite easily by correlation techniques (by noting when a specified threshold is exceeded). The reference time for a specific event occurring was taken as the time of local maximum correlation.

C. Step 3. Signaturing

The ACF representation of an underwater machine sound is shown in Fig. 8, and the presence of a 4 cycle periodicity is clearly indicated. It is intended to demonstrate that WVD signaturing of each of the 4 cylinders is possible, such that a particular cylinder firing could be identified (classified) using these signatures. To do the signaturing the sound signal was first segmented with each segment length corresponding to one cylinder firing. The time span in the first instance is most easily found from the ACF, as explained above, although an estimate can

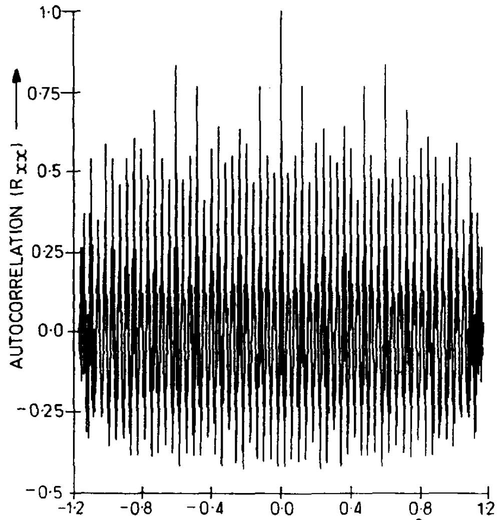


Fig. 10. Autocorrelation of 3 cylinder diesel.

also be obtained by noting the periodicity in the variation of the instantaneous frequency [23]. Wigner–Ville analysis was then performed on 8 consecutive segments of signal (numbered 1–8) from the signal file, with a view to producing signatures for each cylinder.

To produce the WVD plots, a windowing procedure had to be applied [24]. The window length was selected such that there was good frequency resolution as well as good resolution in time; a window length of 395 sample points (5 kHz sampling frequency) was found to give adequate frequency resolution, and 50% overlapping of the windows was used to ensure sufficient detail in the time domain. WVD plots were produced for all eight signal segments (Fig. 11(a)–(h)). Note that in this case the signatures were formed using Wigner–Ville distributions. One could also form signatures using Wigner–Ville spectra, by ensemble averaging the WVD's of many cylinder firings (see Section III).

It was observed that good correspondence existed in the low frequency region of the WVD for segments corresponding to the same cylinder. This result is illustrated by comparing Fig. 11(a) and (e) which correspond to the first and second firings of cylinder 1. Rectangles have been used to highlight the relevant parts of the time-frequency distribution. Similarly Fig. 11(b) and (f), 11(c) and (g), and Fig. 11(d) and (h) represent the first and second firings of cylinders 2, 3, and 4, respectively. It seemed that the cylinders had indeed been signatured. The same sort of signaturing was found to be possible when a 4-cylinder diesel Holden Gemini car was tested [23].

D. Step 4. Classification

For classification of the cylinder to which each of the test signatures belonged, 2D WVD correlation techniques were used (based on the principles in Section III). Table I shows the correlation between segments 5–8 and segments 1–4 (the "References"). The values of highest correlation for a particular segment are underlined, and it is seen that these underlined values correspond to segments which are four segments apart (i.e., along the diagonal), just as they should be for a four-cylinder engine. Thus classification has been successfully achieved.

E. Comparison Between Time Domain Correlations and WVD Correlations

Table II was constructed to compare the performance of WVD signaturing and time domain signaturing. It represents the time correlations between the same eight signal segments shown in Table I. For this table, signal segments 1 to 4 again correspond to the table columns, and signal segments 5 to 8 correspond to the rows. The table elements are the peak correlation values between the various time signals. For each row, the value of highest correlation is underlined; the underlined values are seen to not lie along the diagonal. Thus it appears that a simple time correlation procedure is unable to differentiate between the different cylinders.

Comment: In Section III-A we indicated that time correlations (for white noise) are equivalent to WVD correlations. One reason for the apparent anomaly seen here is that the WVD estimators were formed by performing the 2D cross correlations over the low frequency region shown bounded by rectangles in the time-frequency plots rather than over an infinite frequency range (see Fig. 11(a)–(h)). This calculation corresponds to an effective prefiltering with a low-pass filter. Note that it could, in fact, be obtained in the time domain in this simple case. However, if the signal to be detected has nonstationary spectral characteristics (as in the example presented in Section V) then we apply an effective time-varying prefiltering by multiplying the WVD of the reference signal by a time-varying function $W_h(t, f)$ which preserves (or even enhances) the important features of the signal. Note that this cannot be achieved in the time domain. This approach suffers from problems due to the nonlinearity of the WVD, but as discussed previously, these can be overcome by using the XWVD instead.

For the purposes of comparison, a low-pass filtering was then performed on data in the time domain as described above, and time correlations were again performed between the various signal segments. A sample filtered signal (for firing 2) is shown in Fig. 12). The correlation values thus obtained are shown in Table III. The values of highest correlation (identified by underlining) in Table III do not fall on the diagonal, showing that the time correlations of filtered signal segments are also not reliable in signaturing cylinders. The better performance of the WVD in this case is explained by the fact that the WVD estimators in this case have been formed by a windowing procedure, which has absorbed some of the energy from either side of the actual cylinder boundary to give the signature some information as to the context in which it appears. Thus, while we see that time correlation techniques

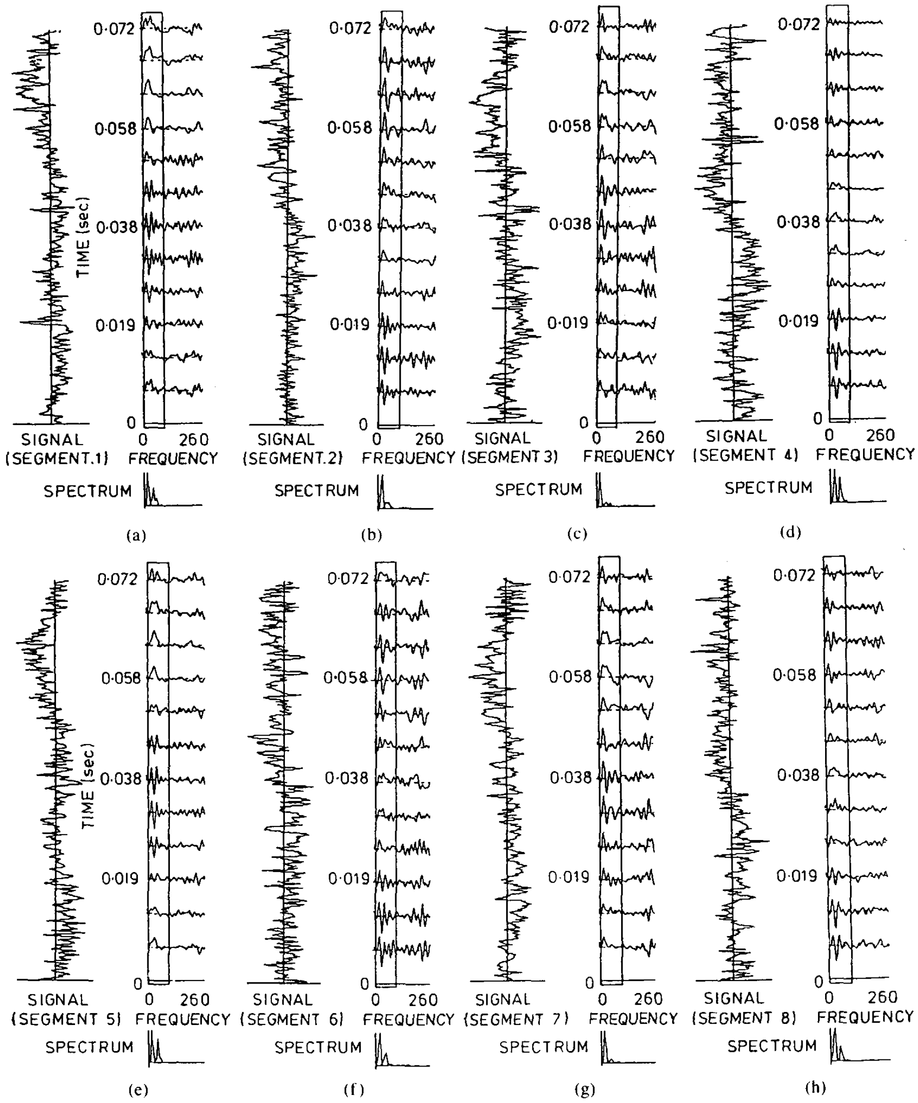


Fig. 11. Low frequency regions of WVD's for 8 consecutive firings of underwater diesel. (a) Firing 1. (b) Firing 2. (c) Firing 3. (d) Firing 4. (e) Firing 5. (f) Firing 6. (g) Firing 7. (h) Firing 8.

TABLE I
WVD CORRELATIONS

	Ref. 1	Ref. 2	Ref. 3	Ref.4
Seg.5	0.889	0.583	0.555	0.536
Seg.6	0.438	0.777	0.467	0.692
Seg.7	0.589	0.535	0.770	0.585
Seg.8	0.326	0.521	0.540	0.835

TABLE II
TIME CORRELATIONS

	Ref. 1	Ref. 2	Ref. 3	Ref.4
Seg.5	0.606	0.356	0.462	0.406
Seg.6	0.314	0.439	0.358	0.449
Seg.7	0.470	0.377	0.558	0.339
Seg.8	0.332	0.443	0.331	0.501

are theoretically equivalent or better in performance than WVD correlations, the potential for preprocessing of the WVD (through window selection, etc.) and postprocessing (time-varying filtering, feature selection, etc.) can make it more a reliable discriminator.

F. Data Averaging

Table I illustrates the separation achievable between different cylinders of a diesel engine with WVD correlations. Given that it is possible to isolate the interval over

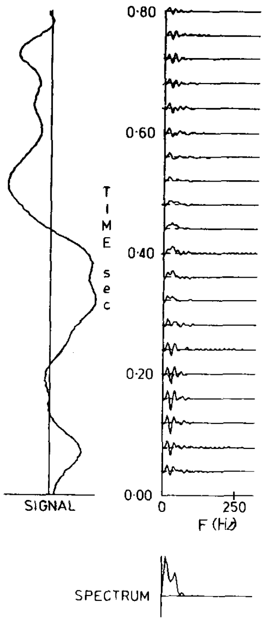


Fig. 12. Filtered WVD of firing 1.

TABLE III
FILTERED TIME CORRELATIONS

	Ref. 1	Ref. 2	Ref. 3	Ref.4
Seg.5	0.843	0.624	0.502	0.758
Seg.6	0.395	0.752	0.612	0.827
Seg.7	0.738	0.684	0.786	0.564
Seg.8	0.644	0.663	0.541	0.811

which a particular cylinder is firing, we decided to average a number of realizations of a particular cylinder firing. The effects of the averaging were evaluated quantitatively by reforming the correlation tables using reference signatures obtained by the averaging procedure. Tables I, IV, and V represent the correlation values between cylinder firings for 3 different time intervals, with each interval having all cylinders fire once. For these tables, segments 1–4 were used as the reference signatures. Tables VI–VIII represent correlation values for the same three intervals (with a small offset) with averaged reference signatures. These correlation values show significantly higher separation between cylinders, indicating quite clearly the importance of averaging these signals. It must be noted that some of this improved separation may be due to using the test signature to form the reference signature (through the averaging process); despite this, genuine improvement does seem to be substantial.

To establish the generality of the techniques, several other known engines were tested to see if their cylinders could also be signatured using this method. The results were affirmative and may be found in [23]. It should be noted, however, that the signatures do change over time.

TABLE IV
WVD CORRELATIONS FOR TIME DISPLACEMENT = 2/CFR

	Ref. 1	Ref. 2	Ref. 3	Ref.4
Seg.9	0.758	0.555	0.542	0.610
Seg.10	0.533	0.744	0.452	0.575
Seg.11	0.729	0.501	0.799	0.573
Seg.12	0.402	0.600	0.459	0.790

TABLE V
WVD CORRELATIONS FOR TIME DISPLACEMENT = 3/CFR

	Ref. 1	Ref. 2	Ref. 3	Ref.4
Seg.13	0.768	0.481	0.475	0.551
Seg.14	0.462	0.561	0.407	0.523
Seg.15	0.601	0.512	0.712	0.502
Seg.16	0.339	0.385	0.514	0.623

TABLE VI
AVERAGED WVD CORRELATIONS FOR TIME DISPLACEMENT = 1/CFR

	Ref. 1	Ref. 2	Ref. 3	Ref.4
Seg. 5	0.857	0.476	0.589	0.608
Seg. 6	0.493	0.836	0.498	0.644
Seg. 7	0.498	0.566	0.892	0.511
Seg. 8	0.573	0.480	0.611	0.917

TABLE VII
AVERAGED WVD CORRELATIONS FOR TIME DISPLACEMENT = 2/CFR

	Ref. 1	Ref. 2	Ref. 3	Ref.4
Seg. 9	0.916	0.464	0.527	0.545
Seg. 10	0.556	0.879	0.485	0.542
Seg. 11	0.577	0.498	0.898	0.525
Seg. 12	0.482	0.587	0.467	0.864

TABLE VIII
AVERAGED WVD CORRELATIONS FOR TIME DISPLACEMENT = 3/CFR

	Ref. 1	Ref. 2	Ref. 3	Ref.4
Seg.13	0.826	0.467	0.525	0.555
Seg.14	0.528	0.677	0.485	0.535
Seg.15	0.431	0.479	0.761	0.478
Seg.16	0.338	0.354	0.442	0.601

G. Use of the XWVD to Signature Cylinders

In Section III it was shown that for optimal detection in the time-frequency domain, the detection statistic should be formed by a 2D correlation between the WVD of the reference signal and the XWVD of the reference and observed signals. This section aims to apply XWVD based detection to the underwater diesel sound, and to compare the results obtained with WVD based detection.

For the initial WVD based scheme implemented in Section IV-D the first firings of cylinders 1–4 were initially assigned as the reference signals. The second set of firings were then assigned to be the observed signals. Table I shows the detection statistics based on these assignments. In Section IV-F the scheme was modified such that the

TABLE IX
XWVD CORRELATIONS

	Ref. 1	Ref. 2	Ref. 3	Ref.4
Obs. 1	0.975	0.594	0.655	0.529
Obs. 2	0.577	0.953	0.513	0.538
Obs. 3	0.609	0.530	0.957	0.544
Obs. 4	0.503	0.595	0.508	0.959

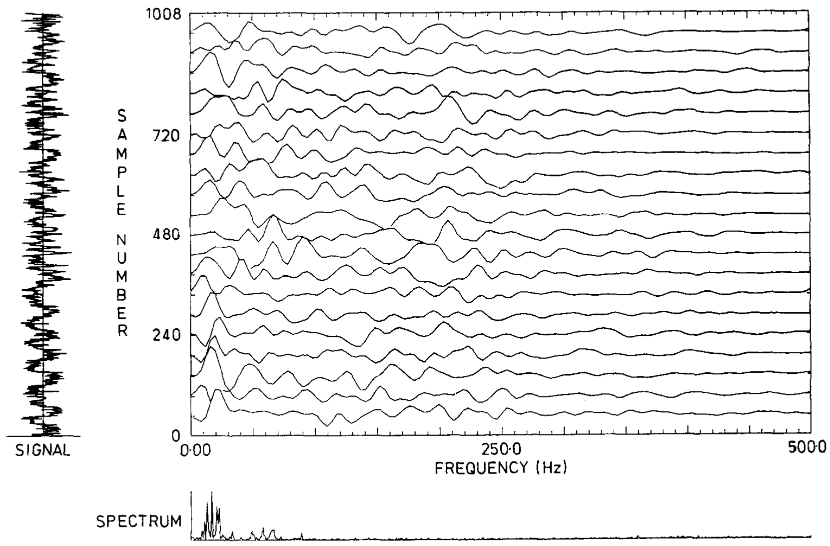


Fig. 13. WVD of transient in 3-dB noise.

reference signal was obtained by an averaging of several firings from the same cylinder. Table VI shows the corresponding detection statistics. There is a general improvement in cylinder separation. Table IX shows the detection statistics when the reference signal is obtained by an averaging, and when the XWVD is used. Significant improvement in separation of the cylinders can be seen, thus supporting the usefulness of the XWVD based scheme.

V. Application 2—Detection of Transient of Unknown Waveshape

Another example of the effectiveness of XWVD based time-frequency correlations incorporating time-varying filtering, is illustrated for real data in Figs. 13–16. Fig. 13 represents the WVD of a real transient (which exhibits a predominantly monocomponent structure) in 3-dB noise. We use the techniques described in Section III-C to first estimate the region in the time-frequency domain where the signal energy is concentrated, and to then adaptively detect the signal. The procedure is as follows.

1) The instantaneous frequency $f_i(t)$ of the signal is first estimated [10].

2) $f_i(t)$ is used to resynthesize a signal $z_x(t)$.

3) The XWVD, $W_{z_x z_r}(t, f)$, is then formed, with its magnitude being shown in Fig. 14.

4) From this magnitude representation, a preliminary bandpass filtering operation is performed, and then thresholding is applied to form a time-frequency window function $W_h(t, f)$ (see Fig. 15). This procedure is expected to enhance the important features of the signal and to eliminate noise.

5) The final estimate for $W_{z_x z_s}(t, f)$ is obtained by multiplying $W_{z_x z_r}(t, f)$ by $W_h(t, f)$, and is shown in Fig. 16.

The value of the detection statistic η_m for H_1 divided by the statistic for H_0, determined according to (13), was found to be 12.9. The corresponding ratio for the case where the signal was estimated using conventional filtering only was found to be 9.7. XWVD based time-frequency correlations, then, perform better in this case.

Simulations were also performed to evaluate the noise performance of the above scheme when applied to a frequency modulated signal with the following parameters: duration: 0.64 s, sampling frequency: 200 Hz, frequency law: linear 15–96 Hz, amplitude envelope: Gaussian rise with standard deviation = 0.1 s and Gaussian fall with standard deviation = 0.56 s, sample length: 128. The performance of the XWVD scheme and the energy detector are shown plotted in Fig. 17. It is seen that the XWVD scheme performs better at all SNR. A different approach to a similar problem was provided in [25]. There the authors used unsupervised weighted maximum likelihood

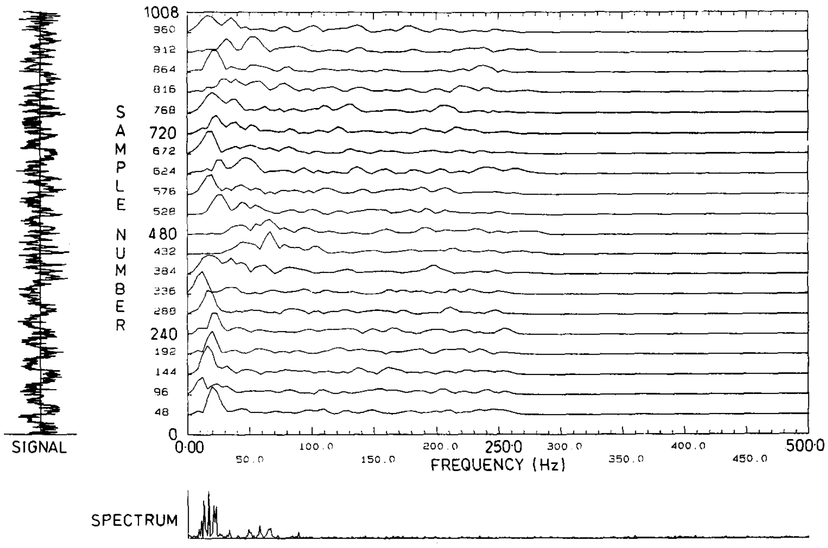


Fig. 14. XWVD (mag.) of transient in 3-dB noise.

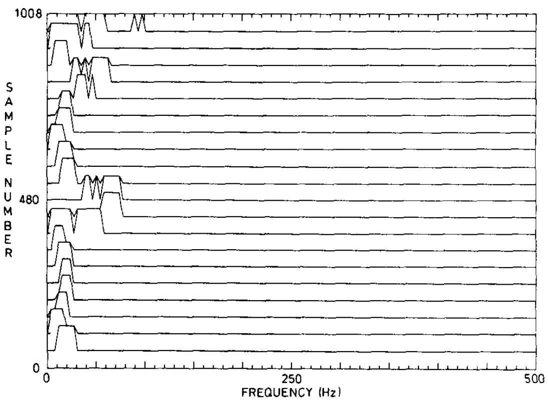


Fig. 15. Window function.

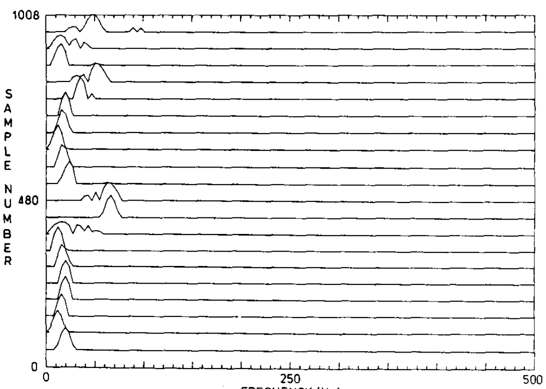


Fig. 16. Windowed XWVD of transient.

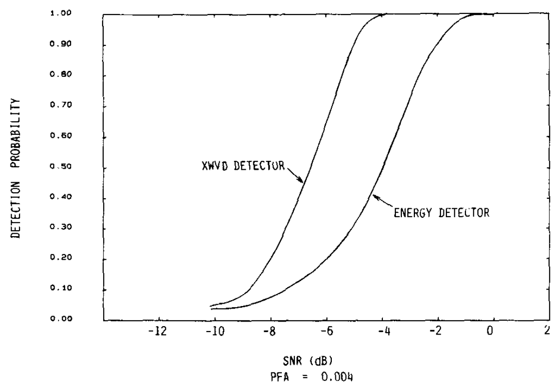


Fig. 17. Performance of the XWVD scheme and the energy detector.

clustering in the Wigner–Ville plane to track and detect mono- and multicomponent frequency modulated signals.

VI. Conclusion

A number of schemes for transient detection have been reviewed, with emphasis on WVD and XWVD based schemes. For the case of white noise, detection schemes using the XWVD are equivalent to the familiar optimal matched filter, while WVD based detection schemes are roughly equivalent for high SNR but degrade in performance as the SNR decreases.

The WVD and XWVD detection schemes were applied to the signaturing of individual engine cylinders. An engine sound signal was segmented into intervals corresponding to one cylinder firing, and time-frequency analysis was performed on each of the segments. It was found that a consistent time-frequency pattern was produced for each particular cylinder. Using one set of time-frequency patterns from each of the cylinders as the reference signatures, a methodology to detect cylinder firing events and to identify the cylinders to which they belonged was described. Correlation techniques were used for the detection, while WVD based techniques were used for the identification. Using the WVD detection scheme and extracting only part of the time-frequency plane, it was possible to correctly identify the particular cylinder, while with the simple time-correlation scheme it was not possible to do so. When the XWVD scheme was used the cylinder could be identified with even greater confidence than with the WVD. The XWVD has also been applied to detection of a nonstationary signal with unknown waveshape and has been shown to perform better than the energy detector.

Acknowledgment

The authors would like to thank R. Barrett and D. McMahon, DSTO, for helpful discussions, and G. Mountford, DSTO, for administrative support.

References

[1] F. B. Tuteur, "On the detection of broad-band targets on noise of uncertain level," *IEEE Trans. Commun. Technol.*, vol. 15, no. 1, 1962.

[2] M. J. Levin, "Power spectrum parameter estimation," *IEEE Trans. Inform. Theory*, vol. IT-11, pp. 100-107, 1965.

[3] B. Porat and B. Friedlander, "Adaptive detection of transient signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, no. 6, Dec. 1986.

[4] C. H. Chen, "Recognition of underwater transient patterns," *Pattern Recognition*, vol. 18, no. 6, pp. 485-490, 1985.

[5] U. Appel and A. V. Brandt, "Adaptive sequential segmentation of piecewise stationary time series," *Inform. Sci.*, vol. 29, pp. 27-56, 1983.

[6] K. Lashkari, B. Friedlander, J. Abel, and B. McQuiston, "Classification of transient signals," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing* (New York), 1988, pp. 2689-2692.

[7] R. A. Altes, "Detection, estimation, and classification with spectrograms," *J. Acoust. Soc. Amer.*, vol. 67, no. 4, pp. 1232-1246, 1980.

[8] B. Friedlander and B. Porat, "Detection of transients by the Gabor representation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 2, pp. 169-180, Feb. 1989.

[9] L. Cohen, "Generalized phase-space distribution functions," *J. Math. Phys.*, vol. 7, no. 5, pp. 781-786, 1966.

[10] B. Boashash, "Time-frequency signal analysis," in *Advances in Spectrum Estimation and Array Processing* (Prentice-Hall Signal Processing Series), S. Haykin, Ed. Englewood Cliffs, NJ: Prentice-Hall, 1990.

[11] B. Bouachache and F. Rodriguez, "Recognition of time-varying signals in the time-frequency domain by means of the Wigner distribution," in *Proc. ICASSP 84*, 1984, pp. 22.5.1-22.5.4.

[12] B. V. K. V. Kumar and C. W. Carroll, "Performance of Wigner distribution function based detection methods," *Opt. Eng.*, vol. 23, no. 6, pp. 732-737, Nov./Dec. 1984.

[13] B. V. K. V. Kumar and C. W. Carroll, "Effects of sampling on signal detection using the cross Wigner distribution function," *Appl. Opt.*, vol. 23, pp. 4090-4094, 1984.

[14] L. Cohen, "On a fundamental property of the Wigner distribution," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 35, no. 4, pp. 559-561, 1987.

[15] S. M. Kay and G. F. Boudreaux-Bartels, "On the optimality of the Wigner distribution for detection," in *Proc. IEEE Conf. Acoust., Speech, Signal Processing* (Tampa, FL), 1985, pp. 1263-1265.

[16] P. Flandrin, "A time-frequency formulation of locally optimum detection," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, no. 9, pp. 1377-1384, Sept. 1988.

[17] B. Boashash, "Note on the use of the Wigner distribution for time-frequency signal analysis," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 1518-1521, Sept. 1988.

[18] J. E. Moyal, "Quantum mechanics as a statistical theory," *Proc. Cambridge Phil. Soc.*, vol. 45, pp. 99-132, 1949.

[19] W. Li, "Wigner distribution method equivalent to dechirp method for detecting a chirp signal," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 35, no. 8, Aug. 1987.

[20] B. Boashash and P. O'Shea, "Time-frequency analysis applied to signaturing of underwater acoustic signals," in *Proc. ICASSP 88* (New York), Apr. 1988, pp. 2817-2820.

[21] T. A. C. M. Claasen and W. F. G. Mecklenbrauker, "The Wigner distribution—A tool for time-frequency signal analysis. Part II: Discrete-time signals," *Philips J. Res.*, vol. 35, pp. 276-300, 1980.

[22] S. L. Marple, *Digital Spectral Analysis with Applications.* Englewood Cliffs, NJ: Prentice-Hall, 1987.

[23] B. Boashash and P. O'Shea, "Time-frequency analysis applied to underwater acoustics—Final report on DSTO Project No. CDS 10/87-DST 86/726," Dep. Elec. Eng. Rep. EE88/7, Univ. of Queensland, Australia, 1988.

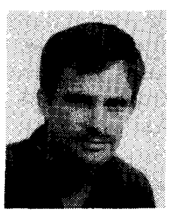
[24] B. Boashash and P. Black, "An efficient real-time implementation of the Wigner-Ville distribution," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 35, no. 11, pp. 1611-1618, Nov. 1987.

[25] F. S. Cohen, G. F. Boudreaux-Bartels, and S. Kadambe, "Tracking of unknown nonstationary chirp signals using unsupervised clustering in the Wigner distribution space," in *Proc. IEEE Conf. Acoust., Speech, Signal Processing* (New York), 1988, pp. 2180-2183.

Boualem Boashash (S'82-M'82-SM'89) was born in Algeria in 1953. He obtained the Diplome d'ingenieur-Physique-Electronique degree from the ICPI University of Lyon in 1978, the M.S. degree in 1979, and the Ph.D. degree (Docteur-Ingenieur) in May 1982, both from the Institut National Polytechnique de Grenoble, Grenoble, France.

In 1979 he joined ELF-Aquitaine Geophysical Research Center, Pau, France, where he was a Research Engineer in the signal processing group. In May 1982 he joined the Institut National des Sciences Appliquees de Lyon, Lyon, France, where he was a Maitre-Assistant associe. In January 1984, he joined the Electrical Engineering Department of the University of Queensland, Australia, as Lecturer, became Senior Lecturer in 1986, and Reader in 1989. In 1990, he joined the Graduate School of Science and Technology, Bond University, as Adjunct Professor of Signal Processing. His research interests are time-frequency signal analysis, spectral estimation, signal detection and classification, as well as real-time signal processing.

Dr. Boashash is the Chairman of the International Symposium on Signal Processing and its Applications, organized regularly in Australia since 1987.



Peter O'Shea (S'88) received the B.E. degree and the Dip. Ed. degree from the University of Queensland, Australia, in 1978 and 1983, respectively. He is currently working towards the Ph.D. degree.

He has worked for the Overseas Telecommunications Commission in Australia, and at a number of educational institutions. From 1987 to 1989 he was employed as a Research Assistant in the CRISSP Laboratory, University of Queensland. His research interests include time-frequency analysis and detection and estimation theory.