BLIND QUADRATIC AND TIME-FREQUENCY BASED DETECTORS FROM TRAINING DATA

Douglas L. Jones and Akbar M. Sayeed

Coordinated Science Laboratory
University of Illinois, Urbana, IL 61801

ABSTRACT

Time-frequency based methods, particularly quadratic (Cohen's-class) representations, are often considered for detection in applications ranging from sonar to machine monitoring. We propose a method of obtaining near-optimal quadratic detectors directly from training data using Fisher's optimal linear discriminant to design a quadratic detector. This detector is optimal in terms of Fisher's scatter criterion as applied to the quadratic outer product of the data vector, and in early simulations appears to closely approximate the true optimal quadratic detector. By relating this quadratic detector to an equivalent operation on the Wigner distribution of a signal, we derive near-optimal time-frequency detectors. A simple example demonstrates the excellent performance of the method.

1. INTRODUCTION

Time-frequency methods have long been used for detection in applications such as sonar and radar. More recently, extensive interest has arisen in time-frequency based detection for new applications as diverse as machine fault detection, communications signal detection, and heart attack risk assessment from ECGs [1] [2]. Time-frequency methods are of interest because of the nonstationary nature of these signals or the interference. However, unlike the classical matched filter theory applied in radar and sonar, the methods developed for these new applications are *ad hoc*, with no assurance that the time-frequency methods employed even approach the best possible performance.

Quadratic detection theory has recently been exploited to derive optimal quadratic time-frequency detectors [3]. Unfortunately, these methods require substantial knowledge of certain statistics of the signals and the interference. In many applications, such as machine fault detection and ECG analysis, the signals are complex and poorly understood, the distinguishing features are unknown, and the collection of the necessary statistics may not be possible. However, the collection of substantial amounts of labelled training data is often feasible. A method based on deriving time-frequency detectors directly from training data offers the best and, perhaps, only hope for optimal detection in many of these important new applications.

In this paper, we develop such a method by applying Fisher's optimal linear discriminant to design a quadratic detector from training data. Relating this quadratic detector to an equivalent operation on the Wigner distribution or the ambiguity function of a signal yields a time-frequency detector. If the training data represents the signal/fault-present case for one particular time-frequency location, we derive under certain circumstances an equivalent detector for components at unknown time-frequency locations. A simple example illustrates the near-optimal performance of the method.

2. QUADRATIC TIME-FREQUENCY ANALYSIS

Most commonly used time-frequency representations (TFRs) are members of Cohen's class, which is the class of quadratic time-frequency representations. All TFRs in Cohen's class can be represented as filtered versions of the Wigner distribution (WD)

$$P(t,\omega) = WD_x(t,\omega) * \phi(t,\omega), \tag{1}$$

where $\phi(t,\omega)$ is the kernel in the time-frequency domain. The WD is a linear transformation of the instantaneous correlation function

$$R(t,\tau) = x(t+\tau/2)x^*(t-\tau/2), \tag{2}$$

where the Wigner distribution is defined as

$$WD_x(t,\omega) = \int_{-\infty}^{\infty} R(t,\tau)e^{-j\omega\tau}d\tau. \tag{3}$$

This work was supported by the Joint Services Electronics Program under Grant No. N00014-90-J-1270, the Schlumberger Foundation, and the Center for Research on Applied Signal Processing.

Any quadratic functional, $Q(x)$, of a signal can be generated via an inner product,

$$Q(x) = (A(t,\tau), R(t,\tau)), \tag{4}$$

between $A(t,\tau)$ and the instantaneous correlation function. This quadratic functional can equivalently be generated via an inner product with the WD as $Q(x) = (\hat{A}, \mathit{WD})$, where $\hat{A}$ is $A(t,\tau)$ transformed according to (3), or from any other invertible linear operator on $R(t,\tau)$. This includes most common TFRs in Cohen's class. We exploit this equivalence below to derive optimal time-frequency detectors for signals with unknown time-frequency shifts.

Note that in practice one usually processes sampled data, in which case the usual discrete-time equivalents of the above representations are used.

3. BLIND QUADRATIC DETECTION

A substantial body of techniques for the "blind" derivation of detectors/discriminators solely from training data have been developed within the discipline of statistical pattern recognition [4]. Fisher's linear discriminant designs a projection vector, $\mathbf{A}$, that minimizes the "scatter," between two groups of labelled training data, of the linear functional $S(r) = (\mathbf{A}, \mathbf{R})$; that is, it approximately maximizes the difference between the output $S(r)$ for the two hypotheses relative to the sum of the standard deviations of the two groups. The detector is implemented by computing the inner product of the projection vector with the data to be classified and comparing this result to a threshold. The details of Fisher's procedure may be found in [4].

By applying Fisher's method to the instantaneous correlation function, $R(t,\tau)$, a quadratic detector can be designed. The projection vector produced by this procedure corresponds to A in (4) above. Fourier transformation in the variable τ produces an equivalent description $\hat{A}$ in terms of the Wigner distribution. The inner product $Q(r) = (A, R)$ is, furthermore, equivalent to a TFR $P(t,\omega)$ evaluated at $(t = 0, \tau = 0)$ with the kernel $\phi(t,\omega) = \hat{A}(t,\omega)$. Equivalent expressions in terms of any other TFR or the ambiguity function are easily derived. Thus, this approach provides a means of deriving, from training data, an "optimal" quadratic discriminator.

We note that the quadratic discriminator derived with this method is optimal only in the sense of maximizing the "scatter" of the training data. For Gaussian inputs under both hypotheses and certain other conditions, Fisher's method converges asymptotically to the optimal linear detector. However, we note that even under these conditions on the sampled input signal, $x(n)$, the method we propose here does not converge to the optimal quadratic detector, because the data seen by the Fisher design algorithm are quadratic products $(x(n)x^*(n+k))$ of the data and are thus not Gaussian-distributed. Nonetheless, at least in our simulations to date, the proposed method converges to a nearly optimal quadratic detector.

4. BLIND TIME-FREQUENCY DETECTION

For detection of random Gaussian signals with a fixed correlation structure (which may be time-varying), or for detection scenarios with a fixed time-frequency offset of the components of interest, a simple quadratic detector suffices. While such detectors *can* equivalently be implemented as inner products with the Wigner distribution or other quadratic TFRs, there is no particular reason to do so. However, if the time-frequency offset varies, as, for example, in a pulse-doppler radar system, a separate quadratic functional at each time-frequency location is required for optimal detection. In this case, the detection problem becomes a composite hypothesis problem with an unknown, or random, time-frequency offset.

In the most general case, where the statistics of the signal and/or interference change with the unknown time and frequency location, a separate quadratic detector, $Q_{t,f}(r) = (A_{t,f}, R)$ must be designed for each location. As a quadratic time-frequency representation, at a given point (t,f), corresponds to an inner product of the kernel at that point with the WD, the time-varying kernel simply corresponds to a transformed version of $A_{t,f}$:

$$\phi_{t,f}(t,\omega) = \int_{-\infty}^{\infty} A_{t,f}(t,\tau) e^{-j\omega\tau} d\tau. \tag{5}$$

A significant drawback in the general case is that sufficient labelled training data must be available to design the unique kernel for each time-frequency location. However, under certain conditions, the kernel will be time and frequency invariant and can be designed from training data consisting of signal-present training data with only a single time-frequency offset. In the special case where the second and fourth-order statistics of the data under both hypotheses at any time-frequency location are simply equivalently shifted time-frequency-shifted versions of the statistics at other time-frequency locations, the optimal kernels derived at each time-frequency location will be identical (except for time-frequency shifts). An example of such a situation is the case where the noise is white and the signal is merely a time-frequency shifted version of some prototype, such as a radar signal. In this case, the optimal kernel is time-frequency-invariant. Furthermore, training data from only a single time-frequency location is needed to design a kernel, and this kernel can then be applied in a TFR to generate an optimal detector for components at any time-frequency location.

In some important applications, the signal and/or noise statistics may be invariant with time, but not with frequency. In intermediate cases such as these, a different quadratic kernel must be designed for each frequency, but for a given frequency, the kernel will be time invariant. the detector corresponds to a bank of (frequency-varying) quadratic filters.

5. EXAMPLE

The following simplistic example illustrates the proposed method. Figure 1a shows 32 samples of a transient signal $x(n) = ne^{-0.25n}$. Its Wigner distribution is shown in Figure 1b. The method described above is used for the blind design from training data of a detector for this signal in white Gaussian noise. The optimal detector of a deterministic signal in white Gaussian noise is well known to be the matched filter; however, the best purely quadratic detector uses the squared magnitude of the matched filter output. By Moyal's relationship, this corresponds to the inner product of the WD of the signal with that of the noisy data. Thus, the optimal quadratic detection kernel in this example, in the time-frequency plane, is the Wigner distribution of the signal. 10,000 realizations of the signal plus noise, and noise only, were generated to train the kernel as described above. The resulting kernel is shown in Figure 1c, and it closely resembles the WD of the signal, which is known to be the optimal quadratic kernel in this simple example.

The kernel can be applied at other time-frequency locations to detect similar transients with unknown time and frequency shifts. Figure 1d shows a realization with a time and frequency shifted version of the transient; Figure 1e shows the TFR of this representation using the kernel shown in Figure 1c. The presence of the component, at the correct time and frequency location, is apparent.

The classification performance of the method was tested by applying the resulting detector to 1000 realizations each of signal present and signal absent, with the decision threshhold set for equal false alarm and miss probabilities. Using the matched filter, the best possible detector, the error probability was 8.5%. With the optimal quadratic detector (a non-coherent matched filter in this case) the error probability was 11.6%. The performance of the quadratic detector designed using the method developed here was 12.9%, which is very close to the performance of the best possible purely quadratic detector. In an experiment with unknown time and frequency shifts, the error probability was 19.3% for the non-coherent matched filter and 23.0% for the blind quadratic detector. These results conform with theory and demonstrate the ability of the proposed blind training to closely approach the performance of the optimal quadratic method.

6. CONCLUSION

The method developed here for designing quadratic and time-frequency detectors requires only training data; it requires no prior knowledge of the signal or interference characteristics. This is potentially of great benefit in the many applications of current interest, such as machinery monitoring, in which the physical systems are so complex that there is little hope of well modeling (or sometimes even understanding), the phenomena to be characterized. In such applications, blind methods such as these may offer the only realistic hope for obtaining near-optimal detection.

As explained above, even under simple Gaussian assumptions on the data, the proposed detector does not in general converge asymptotically to the exactly optimal quadratic detector. While initial simulations indicate convergence to detectors that are nearly optimal, questions regarding the size of the performance gap under more general conditions, and the development of alternative discriminator design procedures tailored for quadratic inputs, warrant further research.

7. REFERENCES

[1] R. A. Rohrbaugh, "Application of time-frequency analysis to machinery condition assessment", in *Proc. 27th Asilomar Conference on Signals, Systems, and Computers*, November 1993.

[2] J. Fang and L. E. Atlas, "Quadratic detectors for noise-immune detection of modulated signals", in *Proc. 27th Asilomar Conference on Signals, Systems, and Computers*, November 1993.

[3] A. Sayeed and D. Jones, "Optimal quadratic detection using bilinear time-frequency and time-scale representations", in *IEEE Int'l Symp. on Time-Frequency and Time-Scale Analysis*, 1994.

[4] R. Duda and P. Hart, *Pattern Classification and Scene Analysis*, John Wiley & Sons, New York, 1973.

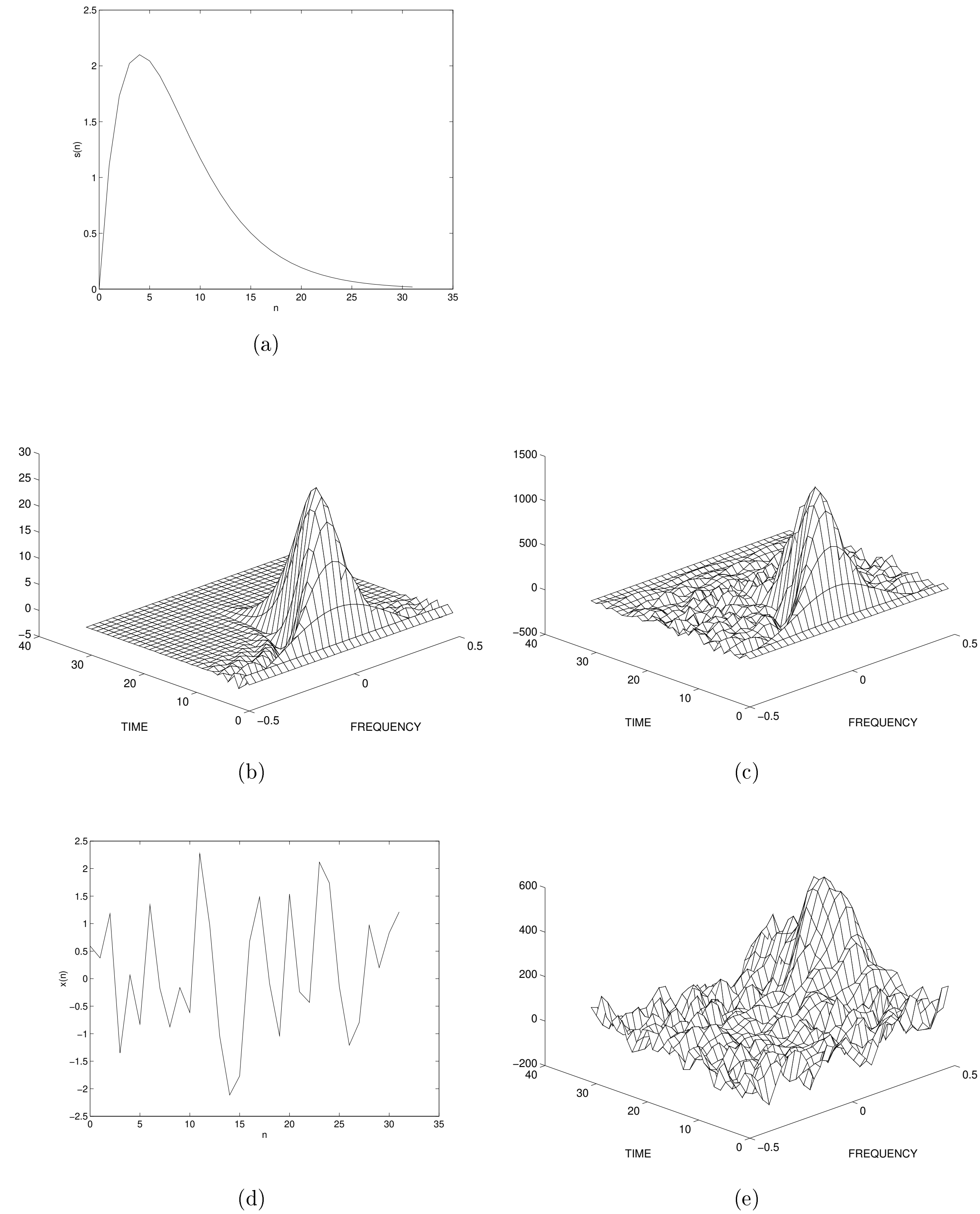

Figure 1: Time-frequency based detection using Fisher's linear discriminant derived from labelled training data. (a) Unknown transient signal to be detected from training data. (b) Wigner distribution of the transient. (c) Wigner distribution of Fisher's kernel. (d) Time-frequency shifted and noisy realization of the transient signal. (e) TFR of the noisy realization based on Fisher's kernel; peak occurs at unknown time-frequency offset.