ENTROPY BASED DETECTION ON THE TIME-FREQUENCY PLANE

Selin Aviyente

Electrical and Computer Engineering
Michigan State University
East Lansing, MI

William J. Williams

Electrical Engineering and Computer Science
University of Michigan
Ann Arbor, MI

ABSTRACT

A comprehensive theory for time-frequency based signal detection has been developed during the past decade. The time-frequency detectors proposed in literature are linear structures operating on the time-frequency representation of the signals and are equivalent to quadratic receivers that are defined in the time domain. In this paper, the concept of entropy based detection on the time-frequency plane is introduced. In recent years, Rényi entropy has been proposed as an effective measure for quantifying signal complexity on the time-frequency plane and some important properties of this measure have been proven. In this paper, a new approach that uses the entropy functional as the test statistic for signal detection is developed. The minimum error detection algorithm is derived and the performance of this new signal detection method is demonstrated through examples.

1. INTRODUCTION

The optimum detection of signals in noise is a well-known problem and has been considered many times in literature [1]. In recent years, there has been approaches to extend signal detection from the time domain to the time-frequency domain since the time-frequency distributions (TFDs) contain more information about nonstationary signals [2, 3, 4, 5]. Cohen's class of TFDs have been extensively used for detection in applications ranging from radar to machine fault diagnosis, due to the need for dealing with nonstationary signals [6]. Most of the time-frequency detectors are linear structures operating in the time-frequency domain and are equivalent to quadratic receivers usually defined in the time domain. It is well-known that the optimal detector of a deterministic signal in white Gaussian noise is the matched filter, and the best purely quadratic detector uses squared magnitude of the matched filter output which can be implemented in the time-frequency domain as the inner product of the Wigner distribution of the observation with the Wigner distribution of the signal. The time-frequency formulation for the optimum detection of Gaussian signals in white Gaussian noise has been proposed by Flandrin [2].

Unfortunately, the design of detectors requires a priori knowledge of signals whereas in real life applications the signals are too complicated and no statistical model is available. Since the collection of labelled signals is often feasible, Jones and Sayeed derived blind time-frequency detectors directly from the training data [3].

For detection of signals in Gaussian noise, matched filter is the optimal detector. However, if the time-frequency offset varies randomly, as for example, in a doppler radar system, a separate quadratic functional at each time-frequency location is required for optimal detection. A significant drawback in the general case is that sufficient training data must be available to design the unique kernel for each time-frequency offset location.

In this paper, we introduce an entropy based detection method. Rényi entropy of time-frequency distributions has been shown to be a robust measure of the complexity of the underlying signal [7, 8]. The detection algorithm is based on the fact that Rényi entropy of a signal plus noise is always less than the entropy of the noise itself. The detection algorithm we propose is invariant to the random time shifts and frequency modulations in the signal, since entropy is a robust measure that is invariant to random shifts in the signal [1]. We follow a data-driven approach as discussed previously in [3, 4]. Unlike the previous approaches, our test statistic cannot be expressed as an inner product of the observation with a reference, and thus does not belong to the class of quadratic detectors.

In Section 2, the background on Rényi entropy and its application to TFDs will be briefly summarized. Section 3 outlines the derivation and the implementation of the minimum probability of error entropy based detection algorithm. In Section 4, we illustrate the performance of this detection algorithm and compare it to the classical matched filtering approach for different types of signals. Finally, Section 5 discusses future extensions of the method presented in this paper.

2. RÉNYI ENTROPY FOR TIME-FREQUENCY DISTRIBUTIONS

The uncertainty of signals are studied indirectly through their time-frequency distributions, which represent the energy distribution of a signal as a function of both time and frequency. A time-frequency distribution, $C(t,\omega)$ from Cohen's class can be expressed as [2] [6]:

$$C(t,\omega) = \frac{1}{4\pi^2}\int\int\int \phi(\theta,\tau)s(u+\frac{\tau}{2})s^*(u-\frac{\tau}{2})e^{j(\theta u-\theta t-\tau\omega)}du\,d\theta\,d\tau \quad (1)$$

where the function $\phi(\theta,\tau)$ is the kernel function and s is the signal. The kernel completely determines the properties of its corresponding TFD. One of the well-known time-frequency distributions is the Wigner distribution given by:

$$W(t,\omega) = \frac{1}{2\pi}\int s(t+\frac{\tau}{2})s^*(t-\frac{\tau}{2})e^{-j\omega\tau}d\tau \quad (2)$$

[1]We concentrate on the class of time-frequency distributions that are time and frequency shift-invariant.

[2]All integrals are from $-\infty$ to ∞ unless otherwise stated.

Wigner distribution a high resolution TFD and therefore is preferred in most time-frequency based detection applications. Some of the most desired properties of TFDs are the energy preservation and the marginals. They are given as follows and are satisfied when $\phi(\theta,0)=\phi(0,\tau)=1 \ \ \forall \tau,\theta$.

$$\begin{aligned}\int\int C(t,\omega)\,dt\,d\omega &= \int |s(t)|^2\,dt=\int |S(\omega)|^2\,d\omega \\ \int C(t,\omega)\,d\omega &= |s(t)|^2 \quad ,\int C(t,\omega)\,dt=|S(\omega)|^2 \end{aligned} \tag{3}$$

The formulas given above evoke an analogy between a TFD and the probability density function (pdf) of a two-dimensional random variable. The main tool in measuring the information content or the uncertainty of a given probability distribution is the entropy function [9]. Williams et al. have extended measures of information from probability theory to the time-frequency plane by treating the time-frequency distributions (TFDs) as density functions [10]. In order to have the TFD behave like a pdf, one needs to normalize it properly, i.e. $C_{normalized}(t,\omega)=\frac{C(t,\omega)}{\int\int C(t,\omega)dtd\omega}$. Another main difference between TFDs and probability density functions is the nonpositivity. Most Cohen's class TFDs are nonpositive and therefore cannot be interpreted strictly as densities of signal energy. Therefore, one should be careful while interpreting the results.

The well-known Shannon entropy when applied to TFDs can be written as:

$$H(C)=-\int\int C(t,\omega)\log_2 C(t,\omega)dtd\omega \tag{4}$$

Since the TFDs are nonpositive in some regions, this definition will not give finite entropy results. For this reason, Rényi entropy has been introduced as a more appropriate way of measuring time-frequency uncertainty [10]. The αth order Rényi entropy is defined as:

$$H_\alpha(C)=\frac{1}{1-\alpha}\log_2\int\int\left(\frac{C(t,\omega)}{\int\int C(u,v)du\,dv}\right)^\alpha dt\,d\omega \tag{5}$$

where $\alpha>0,\alpha\neq 1$. As α goes to 1 Rényi entropy becomes the well-known Shannon entropy functional. It can be shown that for TFDs which are time and frequency shift invariant, Rényi entropy is also shift invariant. Moreover, for the class of scale invariant distributions entropy is invariant to scaling of the signal [7].

3. DERIVATION OF THE DETECTION ALGORITHM

The signal detection in noise problem can be expressed as:

$$\begin{aligned} H_0: y(n) &= v(n) \\ H_1: y(n) &= s(n)+v(n) \end{aligned} \tag{6}$$

where $s(n)$ is the signal and $v(n)$ is white Gaussian noise. The detection will be based on the Rényi entropy of the TFDs. A high entropy value corresponds to a highly random signal, whereas a low entropy value corresponds to a more deterministic signal. Therefore, the decision criterion is:

$$H_\alpha(C_y) \underset{H_1}{\overset{H_0}{\gtrless}} \gamma \tag{7}$$

where γ is the threshold value which will be chosen to minimize the probability of error. This detection criterion is also known as the minimum error detection. The test statistic can be rewritten as

$$\underbrace{\sum_n\sum_k\left(\frac{C_y(n,k)}{\sum_{n'}\sum_{k'}C_y(n',k')}\right)^\alpha}_{X} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{8}$$

where $\eta=2^{(1-\alpha)\gamma}$, and as such it cannot be expressed as a linear functional of the observation $C_y(n,k)$. Therefore, the probability of error is:

$$\begin{aligned} P_e &= P[D_1|H_0]P[H_0]+P[D_0|H_1]P[H_1] \\ &= P_F P[H_0]+P_M P[H_1] \\ P_e &= P[H_\alpha(C_v)<\gamma]P[H_0]+P[H_\alpha(C_{s+v})>\gamma]P[H_1] \end{aligned} \tag{9}$$

where D_1 is the decision of choosing H_1 as the true hypothesis, D_0 is the decision of choosing H_0, P_F and P_M are the probability of false alarm and probability of miss respectively.

The detection algorithm will be derived for a general order of Rényi entropy. The first step in this analysis is to derive the probability density function of the entropy functional under the two hypotheses. For large N and K, large number of time and frequency samples, X in equation 8 can be approximated as a Gaussian with mean m and standard deviation σ by using the central limit theorem argument. Therefore, the pdf for Rényi entropy is derived as follows:

$$\begin{aligned} Y &= \frac{1}{1-\alpha}\log_2\underbrace{\sum_{k=-K}^{K}\sum_{n=-N}^{N}\left(\frac{C(n,k)}{\sum_{n'}\sum_{k'}C(n',k')}\right)^\alpha}_{X} \\ F_Y(y) &= P[Y\leq y]=P\left[\frac{1}{1-\alpha}\log_2 X\leq y\right] \\ &= P\left[X\geq 2^{y(1-\alpha)}\right] \\ f_Y(y) &= \frac{d}{dy}F_Y(y)=\frac{d}{dy}\left[1-F_X(2^{y(1-\alpha)})\right] \\ f_Y(y) &= ((\alpha-1)\ln 2)2^{(1-\alpha)y}\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(2^{(1-\alpha)y}-m)^2}{2\sigma^2}\right) \end{aligned} \tag{10}$$

where the last equality uses the fact that X is distributed as a Gaussian random variable with mean m and standard deviation σ. The threshold for detection is chosen by minimizing P_e, where P_e can be written by substituting equation 10 into equation 9.

$$\begin{aligned} P_e &= P_0\int_{-\infty}^{\gamma}\frac{(\alpha-1)\ln 2}{\sqrt{2\pi}\sigma_0}\exp\left(-\frac{(2^{(1-\alpha)y}-m_0)^2}{2\sigma_0^2}\right)2^{(1-\alpha)y}dy \\ &+ P_1\int_{\gamma}^{\infty}\frac{(\alpha-1)\ln 2}{\sqrt{2\pi}\sigma_1}\exp\left(-\frac{(2^{(1-\alpha)y}-m_1)^2}{2\sigma_1^2}\right)2^{(1-\alpha)y}dy \end{aligned} \tag{11}$$

where P_0 and P_1 are the priors for the two hypotheses, m_0 and m_1 are the means, σ_0 and σ_1 are the standard deviations of the random variable X under noise and signal plus noise hypotheses respectively. To minimize this quantity, we need to solve for γ such that $\frac{\partial P_e}{\partial \gamma} = 0$. The general form for γ is:

$$\begin{aligned} \gamma &= \frac{\log_2 e}{(1-\alpha)} \ln \left[\frac{1}{\sigma_0^2 - \sigma_1^2} (m_1\sigma_0^2 - m_0\sigma_1^2 - \sigma_0\sigma_1\sqrt{A}) \right] \\ A &= m_0^2 + m_1^2 - 2m_0m_1 + 2\sigma_0^2 \ln\left(\frac{\sigma_0 P_1}{\sigma_1 P_0}\right) - 2\sigma_1^2 \ln\left(\frac{\sigma_0 P_1}{\sigma_1 P_0}\right) \end{aligned} \quad (12)$$

The implementation of this detection algorithm uses a training set to estimate the parameters m and σ. The training set consists of an equal number of realizations of H_0 and H_1. From the training set, the mean and the standard deviation of X under the two hypotheses are estimated and used in computing γ. The corresponding probability of error is:

$$P_e = P_0 Q\left(\frac{\eta - m_0}{\sigma_0}\right) + P_1 \left(1 - Q\left(\frac{\eta - m_1}{\sigma_1}\right)\right) \quad (13)$$

where $Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})\, dx$, and $\eta = 2^{(1-\alpha)\gamma}$ where γ, the threshold that yields minimum probability of error, is given by equation 12.

4. SIMULATION RESULTS

In this section, two examples of signal detection will be given using the detection algorithm proposed above. For both examples, a training set consisting of white Gaussian noise and signal plus noise are formed. The mean and the standard deviation of the random variable, X, are estimated from this training set. These parameters are used to evaluate the detection threshold given by equation 12. The derived threshold is used in detecting whether a signal is present or not and a probability of error is computed based on the simulation results. For the purposes of the simulation, order of Rényi entropy, α, will be set to 3 since for that order, the Rényi entropy is well-defined for a large class of signals [7].

Example 1: In this first example, the performances of matched filtering in the time-frequency plane and the entropy based detection method will be compared. Both of the detectors are implemented using a data-driven approach. We consider a randomly shifted gabor logon [3] in white Gaussian noise at SNR$=-6$dB. The training part of the algorithm is done on a set of randomly shifted gabor logons in noise.

For implementing the matched filter, the algorithm derives a reference time-frequency surface such that the discrimination between the two hypotheses, H_0 and H_1, is maximized. When the training is done on a set of gabor logons shifted in time, the reference time-frequency surface comes out to be the average Wigner distribution over all time shifts,

$$C_{ref}(t,\omega) = \frac{1}{T}\int_0^T W_s(t - t_0, \omega)\, dt_0 \quad (14)$$

[3] A gabor logon is a gaussian envelope shifted in time and frequency $x(t) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-t_0)^2}{2\sigma^2}} e^{j\omega_0 t}$.

where t_0 is uniformly distributed between 0 and T.

In the case of the entropy based detection, the location of the gabor logon is not important, since Rényi entropy is invariant to shifts in time and frequency. Therefore, the detector only uses the difference between the entropies of the two hypotheses.

The receiver operating characteristics (ROC) for the two detection algorithms are compared in Figure 1 and it is clear that the entropy based detector performs better than the classical matched filtering. This validates the robustness of our detection algorithm to random shifts. We can quantify the differences in the performances of the matched filter and the entropy based detector by computing the probability of error at the equal error threshold, i.e. $P_F = 1 - P_D$. At this threshold, the probability of error for matched filter is 0.2024 and the probability of error for entropy based detection is 0.078.

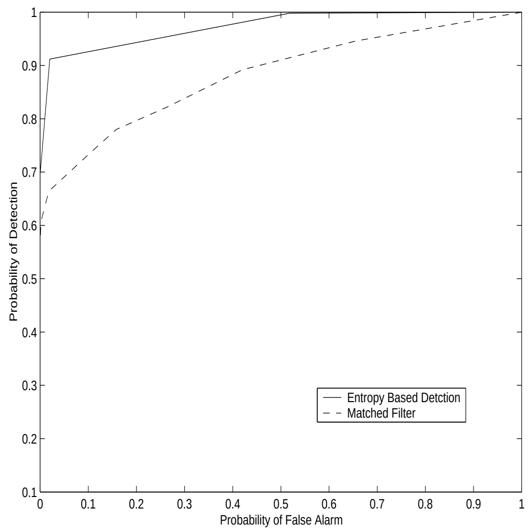

Fig. 1. Comparison of the performances of the matched filter and entropy based detection

Example 2: In this example, we consider the radar backscatter signal. This signal is derived from the Doppler signal produced by a two bladed rotor illuminated by radar [11, 12]. A typical backscatter signal and its Wigner distribution is shown in Figure 2.

We consider a training set consisting of randomly scaled backscatter signals in noise. The threshold for detection is derived using equation 12 and the algorithm is tested on a set of randomly scaled radar backscatter signals in noise. The performance of the algorithm at different SNRs is illustrated in Figure 3. It can be seen that the probability of error is around 0.18 for signal-to-noise ratios as low as -3dB.

This example illustrates the fact that the algorithm performs well for real life signals. It also shows that the detection criterion is robust under a random scaling parameter. This is true since Wigner distribution is scale invariant and for scale invariant distributions Rényi entropy is a scale invariant measure of complexity.

5. CONCLUSIONS

In this paper, we introduced an entropy based detection method on the time-frequency plane. The performance of the entropy based

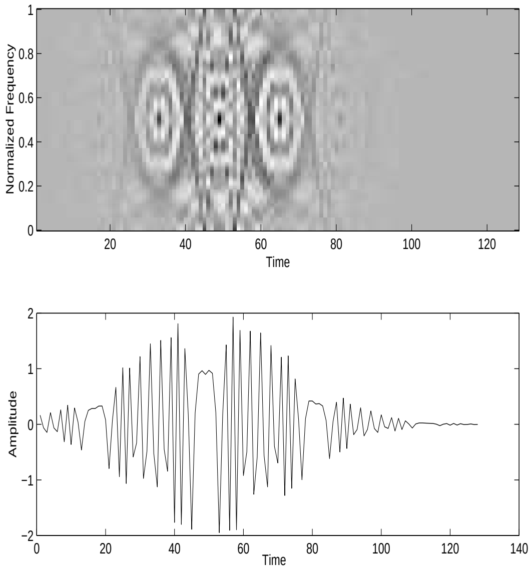

Fig. 2. Radar backscatter signal and its Wigner distribution

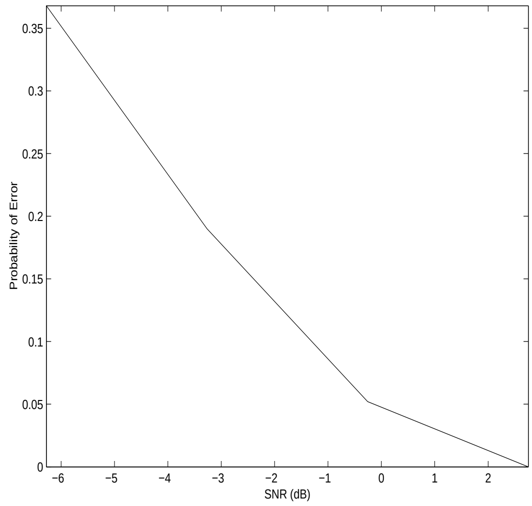

Fig. 3. Probability of error versus SNR for the radar backscatter signal in noise

detection algorithm depends on the difference of the complexities of the signals under the two hypotheses. Therefore, the algorithm works best for the detection of a gabor logon in white noise since it is the most concentrated signal on the time-frequency plane and has the smallest entropy value. The entropy based detection algorithm was tested on signals other than the gabor logon such as the radar backscatter signal in Example 2. The algorithm performs consistently well with other signal types and is superior to classical matched filtering in cases where the signal is randomly shifted in time and frequency. It is possible to extend the algorithm for signals with random parameters other than the time and frequency shift parameters by adding a simple parameter estimation step very much like the standard detection algorithms.

Another important issue for further research is determining the optimal value of α, order of Rényi entropy, to minimize the probability of error. Preliminary results suggest that it is a signal dependent quantity.

6. REFERENCES

[1] H. L. VanTrees, *Detection, Estimation and Modulation Theory, Part I,* Wiley, New York, 1971.

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

[3] D. L. Jones and A. M. Sayeed, "Blind quadratic and time-frequency based detectors from training data," in *Proc. IEEE Int. Conf. Acoustics,Speech and Signal Processing*, 1995, vol. 2, pp. 1033–1036.

[4] C. Richard and R. Lengellé, "Data-driven design and complexity control of time-frequency detectors," *Signal Processing*, vol. 77, no. 1, pp. 37–48, 1999.

[5] A. M. Sayeed and D. L. Jones, "Optimal detection using bilinear time-frequency and time-scale representations," *IEEE Trans. on Signal Processing*, vol. 43, no. 12, pp. 2872–2883, Dec. 1995.

[6] L. Cohen, *Time-Frequency Analysis*, Prentice Hall, New Jersey, 1995.

[7] R. G. Baraniuk, P. Flandrin, A. J. E. M. Janssen, and O. Michel, "Measuring time-frequency information content using the Rényi entropies," *IEEE Trans. on Info. Theory*, vol. 47, no. 4, pp. 1391–1409, May 2001.

[8] S. Aviyente and W. J. Williams, "Information bounds on the time-frequency plane," in *Proc. IEEE Int. Conf. Acoustics,Speech and Signal Processing*, 2001, vol. 6, pp. 3549–3552.

[9] T. M. Cover and J. A. Thomas, *Elements of Information Theory,* Wiley Series in Communications, New York, 1991.

[10] W. J. Williams, M. Brown, and A. Hero, "Uncertainty, information and time-frequency distributions," in *SPIE-Advanced Signal Processing Algorithms*, 1991, vol. 1556, pp. 144–156.

[11] C. J. McCormack, V. V. Liepa, and W. J. Williams, "Time-frequency analysis of radar target backscatter," in *SPIE-Advanced Signal Processing Algorithms*, 1995, vol. 2563, pp. 113–124.

[12] W. J. Williams and E. J. Zalubas, "Invariant classification of time-frequency representations: Applications to doppler target identification," in *Proc. of the US/Australia Joint Workshop on Defence Applications of Signal Processing*, 2001, pp. 278–295.