Cyclostationary Approaches to Signal Detection and Classification in Cognitive Radio

Kyouwoong Kim, Ihsan A. Akbar, Kyung K. Bae, Jung-sun Um*, Chad M. Spooner**, and Jeffrey H. Reed
Virginia Polytechnic Institute and State University
Wireless @ Virginia Tech
Mobile and Potable Radio Research Group
432 Durham Hall, Mail Stop 0350, Blacksburg, VA 24061, USA
*Electronics and Telecommunications Research Institute (ETRI), Daejeon, Korea
**NorthWest Research Associates
{kyouwook, iakbar, kbae, reedjh}@vt.edu, korses@etri.re.kr, cmspooner@nwra.com

***Abstract*— Spectrum awareness is currently one of the most challenging problems in cognitive radio (CR) design. Detection and classification of very low SNR signals with relaxed information on the signal parameters being detected is critical for proper CR functionality as it enables the CR to react and adapt to the changes in its radio environment. In this work, the cycle frequency domain profile (CDP) is used for signal detection and preprocessing for signal classification. Signal features are extracted from CDP using a threshold-test method. For classification, a Hidden Markov Model (HMM) has been used to process extracted signal features due to its robust pattern-matching capability. We also investigate the effects of varied observation length on signal detection and classification. It is found that the CDP-based detector and the HMM-based classifier can detect and classify incoming signals at a range of low SNRs.**

Index Terms - Cognitive Radio, Cyclostationarity, HMMs, Signal Classification

I. INTRODUCTION

Spectrum scarcity is driven mainly by regulatory and license processes for spectrum allocation and not by the fundamental lack of spectrum [1]. Realizing the fact that the licensed spectrum remains unused most of the time, the Federal Communication Commission (FCC) is considering a paradigm shift on spectrum allocation policy towards the adoption of unlicensed, rule-based strategies for certain frequency bands [2].

In this paper, a robust signal detection and pattern matching based signal classification algorithm using cyclostationarity of signals is proposed, which provides the necessary information about the radio environment to the CR to enhance spectral efficiency. Our proposed technique is found to be fairly robust against strong Additive White Gaussian Noise (AWGN). HMMs, trained with the CDP of different signals, are used as a signal classifier due to their promising pattern-matching capabilities.

II. CYCLOSTATIONARY SPECTRAL ANALYSIS

A. Background

A process $x(t)$ is said to be cyclostationary in wide sense if its mean and autocorrelation are periodic with a period T_0; $M_x(t+T_0) = M_x(t)$ and $R_x(t+T_0, u+T_0) = R_x(t,u)$ for all t and u. The relative time difference is of significant importance and the cyclic autocorrelation can be expressed in terms of t and τ as $R_x(t+\tau/2, t-\tau/2)$. This is periodic in t with period T_0 and can be expressed as a Fourier series [3], [4]

$$R_x(t+\tau/2, t-\tau/2) = \sum_{\alpha} R_x^{\alpha}(\tau) e^{j2\pi\alpha t} \tag{1}$$

where $\alpha = m/T_0$ and m is an integer. The Fourier coefficient can be obtained by,

$$R_x^{\alpha}(\tau) \triangleq \lim_{Z\to\infty} \frac{1}{Z} \int_{-Z/2}^{Z/2} R_x(t+\tau/2, t-\tau/2) e^{-j2\pi\alpha t} dt \tag{2}$$

for more than one periodicity. The cyclic Wiener relation states [3] that the spectral correlation function (SCF) can be obtained from the Fourier transform of the cyclic autocorrelation in Eq. 2,

$$S_x^{\alpha}(f) = \int_{-\infty}^{\infty} R_x^{\alpha}(\tau) e^{-j2\pi f\tau}\, d\tau \tag{3}$$

In practical situations, however, the number of observation samples at the sensor is limited. Therefore, the spectral correlation needs to be estimated from a finite set of samples. We use the spectrally smoothed cyclic periodogram method [5]. Let us define the cyclic periodogram by [4] ,[6],

$$S_{x_T}^{\alpha}(t,f) \triangleq \frac{1}{T} X_T(t, f+\alpha/2) X_T^*(t, f-\alpha/2) \tag{4}$$

where X_T is the time-variant Fourier transform defined as follows,

$$X_T(t,f) \triangleq \int_{t-T/2}^{t+T/2} x(u) e^{-j2\pi f u} du \tag{5}$$

The estimated SCF obtained by frequency smoothing of the cyclic periodogram in Eq. 4 is,

$$S_{x_T}^{\alpha}(t,f)_{\Delta f} \triangleq \frac{1}{\Delta f} \int_{f-\Delta f/2}^{f+\Delta f/2} S_{x_T}^{\alpha}(t,v) dv \tag{6}$$

1-4244-0663-3/07/$20.00 ©2007 IEEE

The SCF can be obtained by increasing observation length T and reducing the size of the smoothing window Δf,

$$S_x^\alpha(f) = \lim_{\Delta f \to 0} \lim_{T \to \infty} S_{x_T}^\alpha(t, f)_{\Delta f} \tag{7}$$

From an engineering viewpoint, finding the appropriate averaging sizes in the temporal and spectral domains is important. In this paper, we show the effects of the averaging parameters on the cyclic features.

B. Spectral Coherence and Cycle Frequency Domain Profile

The SCF is a cross-correlation function between frequency components separated by $f + \alpha/2$ and $f - \alpha/2$. It is natural to define the correlation coefficient for the SCF, which is known commonly as spectral coherence (SC). Mathematically, the spectral coherence is defined as [3], [4],

$$C_x^\alpha(f) \triangleq \frac{S_x^\alpha(f)}{[S(f+\alpha/2)S(f-\alpha/2)]^{1/2}} \tag{8}$$

The magnitude of the SC ranges from 0 to 1, but the SC itself is on the (closed) unit disk. It allows us to measure the strength of second-order periodicity contained within a time series in a unified way. In addition, the SC is invariant to the linear transformation of the incoming signal if it does not eliminate the cyclic features at cycle frequency domain. For instance, a filter can reject specific cyclic features at cycle frequencies by changing spectral shape.
The cyclic frequency is inherently a discrete parameter. It is convenient to visually assess the cycle-frequency parameter by evaluating the maximum over spectral frequency f,

$$I(\alpha) \triangleq \max_f |C_x^\alpha(f)| \tag{9}$$

This results in generating the cycle frequency domain profile (CDP) (also called α-domain profile [7]).

C. Cyclostationary Spectral Analysis of Modulated Signals

To evaluate the performance of both detection and classification at low SNR using SC, we used the following simple signal model,

$$y(t) = x(t) + n(t) \tag{10}$$

where $n(t)$ denotes AWGN. To see the detection and classification performance of the proposed technique for various modulation type, the modulation types considered for $x(t)$ include double-side band suppressed carrier AM (DSB-SC AM), BPSK, binary FSK (b-FSK), MSK, and QPSK. We assume that the sampling frequency is twice the incoming signal bandwidth. The cycle-frequency domain profiles for BPSK and QPSK at the SNR of 0dB are shown in Figure 1.

The plots are generated through SC estimation, i.e., time-domain averaging and frequency-domain smoothing are applied to the cyclic periodogram in Eq. 4. For time-domain averaging, we propose a unique method in which multiple observations of the limited time duration T (500 samples are used for simulation) are first obtained and the average value is then obtained using all these observations. This method is easily scalable for the various lengths of time series and allows us to use smaller size FFT for larger time samples which exceeds the given FFT size. For instance, the observation length N = 100 implies the total observation samples of $100 \times T$ and the average value is attained as,

$$\tilde{S}_{x_T}^\alpha(f) \triangleq \frac{1}{N} \sum_{k=1}^{N} S_{x_T}^\alpha(t_k, f) \tag{11}$$

If N increases, the erratic behavior of SC due to the added noise is reduced and we obtain distinct features inherent to the specific signal type at the cycle frequency domain. One of interesting impacts of modulation parameters on the CDP is the roll-off factor of SQuared-root Raised Cosine (SQRC) pulses. The amount of excess bandwidth caused by different roll-off factors in SQRC filtering affects the dominant cyclic features. The cyclic feature can not be detected for the signal which is SQRC filtered with excess bandwidth around 0%. Thus, different modulation and pulse shaping filter result in varying amplitude of cyclic features and they thwart to set up a constant threshold for signal detection for all modulation types. The SNR used in following figures and simulations is defined as the ratio of the signal power to the average noise power within one symbol duration.

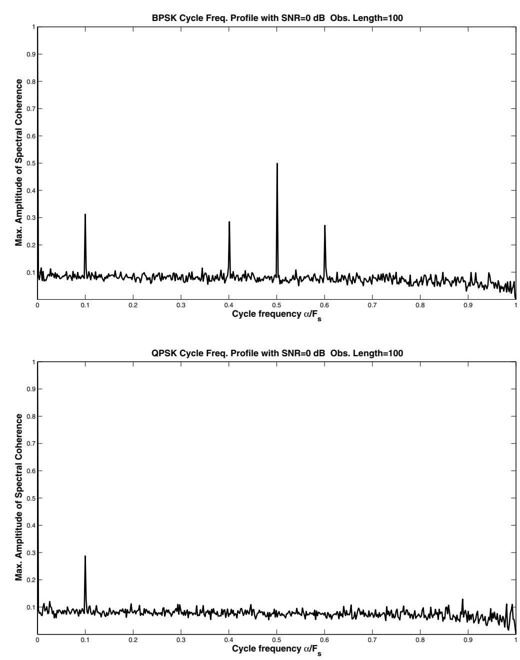

Fig. 1. Cycle Frequency Domain Profile for BPSK (top) and QPSK

III. DETECTION AND FEATURE EXTRACTION

One of the promising characteristics of the cyclic spectral analysis is that the stationary noise added to the signal having noise uncertainty or variability can be suppressed almost completely. This means that we can observe distinct features even at low SNRs with noise uncertainty if we increase the number of blocks used for averaging. We assumed that the rough information of bandwidth is known to the signal sensor. The signal sensor does not know modulation parameters such as carrier frequency, symbol rate, and pulse shaping. In that case, the signal sensor must scan all cycle frequencies. We introduce the crest factor (CF) for signal detection and feature extraction by exploiting CDP shape. The crest factor of a waveform is equal to the peak amplitude of a waveform divided by the root mean square (RMS). The CF is a dimensionless quantity. Signal detection using CDP is finding the highest peak in CDP. When the highest peak location is known, this boils down to optimal single cycle detector [8].

For signal detection, the threshold C_{TH} is first calculated when no signal is present, i.e., when $x(t) = n(t)$ in Eq. 10,

$$C_{TH} = \max(I(\alpha)) / \sqrt{\left(\sum_{\alpha=0}^{N} I^2(\alpha)\right) / N} \tag{12}$$

C_{TH} is a random value due to the random noise $n(t)$. False alarm rate is obtained by estimating probability density function (PDF) of C_{TH}. The estimated PDF of C_{TH} can be evaluated by plotting a histogram of C_{TH}. To test signal existence in additive white Gaussian noise (AWGN), following binary hypothesis testing is performed.

$$\begin{aligned} H_0 &: x(t) = n(t) \\ H_1 &: x(t) = s(t) + n(t) \end{aligned} \tag{13}$$

The proposed threshold-based signal detection test can be performed as follows

$$\begin{aligned} C_I \leq C_{TH} &: \text{Declare } \mathrm{H}_0 \\ C_I > C_{TH} &: \text{Declare } \mathrm{H}_1 \end{aligned} \tag{14}$$

For feature extraction, all CDP peaks greater than C_{TH} are encoded one and the others are encoded zero. This generated binary feature vector is fed into the HMM signal classifier.

IV. HMM AS SIGNAL CLASSIFIER

A hidden Markov process (HMP) [9] is a doubly stochastic process in which the generation of observation symbols depends on the emission properties of the states. Therefore, a state can generate more than one observation symbol and the state sequence is not directly observable, given the observation sequence.

Mathematically, an HMP can be defined as the pair $\{X_t, Y_t;\ t \in \mathbf{N}\}$ of stochastic processes defined on the probability space. Here (Ω, F, P) denotes the hidden state sequence and Y_t is the observation sequence. The finite set (X, Y) is said to be a stationary finite state system (SFSS) if the following conditions are met:

- (X_t, Y_t) are jointly stationary,
- $\Pr(Y_{t+1} = y_{t+1}, X_{t+1} = x_{t+1} | Y_1^t = y_1^t, X_1^t = x_1^t) = \Pr(Y_{t+1} = y_{t+1}, X_{t+1} = x_{t+1} | X_t = x_t)$

The processes X_t and Y_t are called the state and the output of the SFSS respectively. The mathematical model that can generate such a HMP is called a hidden Markov model (HMM). A HMM [10] is a finite state machine in which the observation sequence is a probabilistic function of states. It differs from Markov chains (MCs) [11] where the observation sequences is a deterministic function of states. In case of MCs, the states are directly observable from the observation sequence while the state sequence is not directly observable for the observed data and is hidden in the case of HMMs. HMMs were first introduced as a pattern recognition tool by Rabiner [12] in the beginning of the 1970's. Ever since, these models are used in many areas of sciences and engineering because of their strong mathematical structure and theoretical basis.

A discrete time HMM with N states and M symbols consists of an N by N state transition matrix $\mathbf{P}$ that defines the probability of transitioning from one state to another or itself, an N by M output symbol probability matrix $\mathbf{B}$ that gives the probability of generating different output symbols while being in a particular state, and an N dimensional vector called the initial state probability vector that gives the probability of being in a particular state at the start of the process. A hidden Markov model is denoted as $\zeta = \{\mathbf{P}, \mathbf{B}, \pi\}$. These parameters can be estimated using the Baum-Welch Algorithm (BWA), which is basically a derived form of the Expectation-Maximization (EM) algorithm for HMMs. Due to the need for an online estimation in real world applications, we use a modified version of the BWA, called as the Forward-only BWA (FO-BWA) that can estimate HMM parameters on the fly. Details about FO-BWA can be found from [13].

For the case of binary sequences, the probability of generating the observation sequence given the model can be written mathematically as

$$P(y_1^T|\zeta) = \pi\mathbf{B}(y_1)\ \mathbf{PB}(y_2)\ \mathbf{P} \ldots\ \mathbf{PB}(y_T)\mathbf{1}^t \tag{15}$$

Here $\mathbf{B}(y_k)$ with $k = 1, 2, \ldots, T$ denotes the probability of generating symbol from different states. Because of the significantly long data size, we use the logarithm of $P(y_1^T|\zeta)$, usually known as log-likelihood.

V. SIGNAL CLASSIFICATION

If the CDP based detector declares that a signal exists, then this signal goes through the signal classification stage. For training purposes, ideal binary feature vectors are generated using CDPs for various signal types. The feature vectors are fed into the HMM for learning process that uses the Baum-Welch algorithm. The Baum-Welch algorithm produces Hidden Markov models, $\zeta = \{\mathbf{P}, \mathbf{B}, \pi\}$, based on each training sequence (signal type). After successful training, the unknown incoming signal is used to find its likelihood using each HMM generated in the training phase. The likelihood values hence generated is compared with the likelihood of the original sequence and the closest match is selected as the signal type. A simplified block diagram of signal classification is shown in Fig. 2.

VI. COMPUTER SIMULATION RESULT

A. Signal Detection

We analyze the relationship between observation length and the SNR of the incoming signal for signal detection with 10% false alarm. Different values of observation lengths are plotted against incoming SNRs for signal detection in Fig. 3.

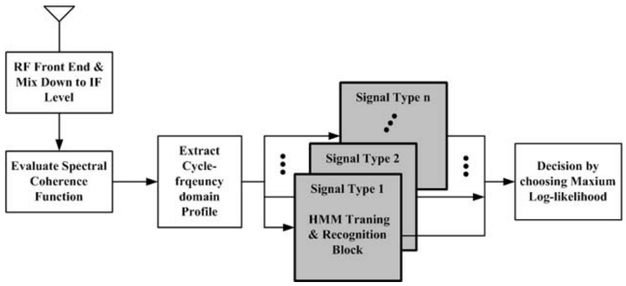

Fig. 2. Cycle-HMM Recognition Block Diagram

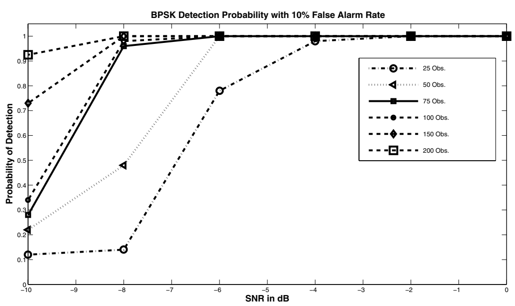

Fig. 3. BPSK Detection Probability of 10% False Alarm

B. Signal Classification

Monte Carlo simulations were performed for signal classification. The HMMs in Fig. 2 were trained with ideal feature vectors for each signal type. Different incoming signals with SNR of -3dB are observed with varying observation length to obtain the percentage of successful classification. The result is summarized in Fig. 4. Note that the percentage of correct signal classification (for each signal type) reach 100 when we increase the observation length to 300 blocks. The CDP for AM and QPSK have close similarity. However, QPSK CDP is more susceptible to the noise than AM, hence the classification performance is not good in low SNR. Usually, cyclic feature related to the pure sine wave is stronger than the feature caused by symbol rate.

VII. CONCLUSIONS

In this paper we proposed and investigated techniques for detection and classification of radio signals in a cognitive radio (CR) environment. Simulation results show that we can detect the incoming signals, even at very low SNR, if the number of observation blocks is sufficiently large. One of the method's advantages is that it does not require any

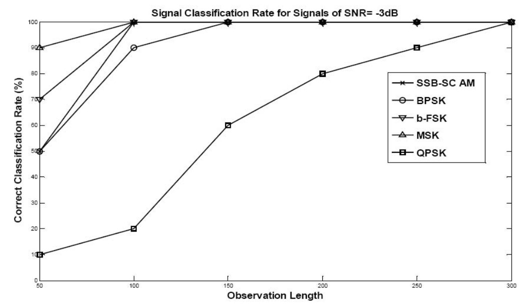

Fig. 4. Percentage of Successful Classification with -3dB Incoming Signals

a priori knowledge of the transmitting signal except rough information on signal bandwidth. It allows the CR device, which tries to access a specific channel in an opportunistic manner, to blindly detect active or licensed user signals in that channel. Signal classification can be performed with high accuracy if the observation length is sufficiently long. In a CR application, the training sequence can be retrieved from a database which is maintained in the sensor itself, in the local spectrum management agency, or in base stations. In this work, the signal detection and classification were performed in stationary noise environment such as an AWGN channel. The proposed methods should be tested in a fading channel environment with noise uncertainty and for more sophisticated modulation schemes.

ACKNOWLEDGMENT

The authors thank ETRI for providing technical as well as financial support that made this work possible.

REFERENCES

[1] M. A. McHenry, "NSF Spectrum Occupancy Measurements Projects Summary," Shared Spectrum Company Report, Aug. 2005.
[2] FCC 03-322, "NPRM - Facilitating Opportunities for Flexible, Efficient, and Reliable Spectrum Use Employing Cognitive Radio Technologies," FCC, Dec. 2003.
[3] W. A. Gardner, Introduction to Random Process with Applications to Signals and Systems, MacMillan, 1986.
[4] W. A. Gardner, Statistical Spectral Analysis A Nonprobabilistic Theory, Prentice Hall, 1988.
[5] R. S. Roberts, "Computationally Efficient Algorithms for Cyclic Spectral Analysis," IEEE Signal Processing Magazine, Apr. 1991.
[6] W. M. Gardner, "Measurement of Spectral Correlation," IEEE Transaction on Acoustics, Speech, and Signal Processing, vol. ASSP-34, No. 5, Oct. 1986
[7] A. Fehske, J. Gaeddert and J. H. Reed, "A New Approach to Signal Classification Using Spectral Correlation and Neural Networks," IEEE International Symposium on New Frontiers in DySpan, Nov. 2005.
[8] W. M. Gardner and C. M. Spooner, "Signal Interception: Performance Advantages of Cycle-Feature Detectors," IEEE Transactions on Communications, vol. 40, no.1, Jan. 1992.
[9] Y. Ephraim and N. Merhav, "Hidden Markov Processes," IEEE Transactions on Information Theory, vol. 48, no. 6, July 2002.
[10] L. R. Rabiner and B. H. Juang, "An introduction to Hidden Markov Models," IEEE ASSP Magazine, January 1986.
[11] S. M. Ross, Introduction to Probability Models, Academic Press, Inc., 1993.
[12] L. Rabiner and B. H. Jaung, Fundamentals of Speech Recognition, Prentice Hall, Englewood Cliffs, New Jersey, 1993.
[13] W. Turin, Digital Transmission Systems: Performance Analysis and Modeling, McGraw-Hill, 1998.