ROBUSTIFICATION OF CYCLOSTATIONARY ARRAY PROCESSING TECHNIQUES

Paul Burns and Dr. Zhi Ding
Department of Electrical Engineering, Auburn University, Auburn, AL 36849-5201

Abstract
The problem of exploiting cyclostationary statistical information for the purpose of array processing is addressed. Techniques for exploiting second order periodic information are proposed for the enhancement of cyclic-MUSIC and Self COherent REstoration (SCORE) algorithms. The robustification of both cyclic MUSIC and SCORE can be accomplished by increasing the amount of cyclic information used. Simulation results are presented which show the performance improvement by the modified algorithms.

1 Introduction

Recent development in cyclostationary signal processing by Gardner et al has resulted in two effective statistical array processing methods. The first is a subspace direction of arrival (DOA) estimation approach called cyclic-MUSIC [1]. The other is a blind spatio-temporal filtering algorithm called SCORE (Self- COherent REstoral) [2]. SCORE has been shown to give the optimum solution for the signal extraction problem using cyclostationary statistics [2]. Because both of these methods employ the cyclic correlation matrix, they both can potentially be improved by exploiting many cyclic correlation matrices (evaluated at various lags, τ). This paper addresses the issues of how to effectively use the information given in more than one cyclic correlation matrix with cyclic-MUSIC and SCORE and how to robustify the algorithms and improve their overall performance.

2 Problem Description

Consider a receiver platform consisting of n monopole antennas physically aligned in a row. Each antenna receives a signal, $x_i(t)$, multiplies it by a factor w_i^*, and sums it with other signals, producing $y(t)$. The relationship between the inputs and the array output is conveniently formulated in vector form as

$$y(t) = \mathbf{w}^H \mathbf{x}(t). \tag{1}$$

For the subspace DOA estimation algorithms considered here, the data $\{x_i(t)\}$ are used to form direction estimates. For spatial filtering approaches the signal $\mathbf{x}(t)$ is processed by $\mathbf{w}$ to yield a desirable $y(t)$. The typical assumption that all SOI (signals of interest) and interferences are narrow-band plane waves will be made here. For a plane wave impinging at an angle of θ, the electrical phase shift, ϕ, is $\phi = \frac{d}{\lambda}\sin(\theta)$ where λ is the wavelength of the signal carrier and d is the distance between uniform antennas. Hence, we can express the received signal as

$$\mathbf{x}(t) = \sum_{i=1}^{L} \mathbf{a}(\theta_i) s_i(t) + \mathbf{n}(t) \tag{2}$$

where $s_i(t)$ is a signal impinging from an angle θ_i, L is the total number of plane waves impinging on the array, and $\mathbf{n}(t)$ is receiver noise. $\mathbf{a}(\theta)$ is a direction "steering" vector of the form $\mathbf{a}(\theta) = [1\ e^{j\phi} \cdots e^{j(n-1)\phi}]^T$ which models the array response to narrow-band plane waves.

2.1 Conventional MUSIC

Before cyclic-MUSIC is described, the conventional MUSIC algorithm will be briefly explained here.

Consider once again the signal $\mathbf{x}(t)$ in equation (2). Suppose there are n antennas in the array and L signals in the environment. The purpose of MUSIC is to estimate the DOA of each of the L signals. In addition to the above assumptions, it will be assumed that the sensor noise is white and zero mean and that the L signals are uncorrelated and arrive at the array from L distinct angles.

Under these assumptions, $\mathbf{x}(t)$ is expressed as in equation (2),

$$\mathbf{x}(t) = A(\theta)\mathbf{s}(t) + \mathbf{n}(t) \tag{3}$$

and its autocovariance matrix is

$$R_{xx} = A(\theta) R_{ss} A^H(\theta) + \sigma_n^2 I_{n\times n} \tag{4}$$

where $A(\theta)$ is a $n \times L$ matrix with i^{th} column equal to $\mathbf{a}(\theta_i)$. σ_n^2 is the variance of the noise. The eigenvectors corresponding to the L largest eigenvalues provide a basis for the signal subspace. Those corresponding to the $(n-L)$ smallest eigenvalues are a basis for the noise subspace. Because R_{xx} is positive definite, those two subspaces are orthogonal.

The principle behind MUSIC is to obtain an estimate of the noise subspace, V_n, from an estimate of R_{xx}. The DOA estimates of impinging signals can then be determined by finding the maxima of

$$F(\theta) = \frac{1}{||V_n \mathbf{a}(\theta)||^2}. \tag{5}$$

Those maximizing θ's are the DOA estimates.

0-7803-2431-5/95 $4.00

2.2 Cyclic MUSIC

Cyclic-MUSIC differs from conventional MUSIC only in that the null space of the *cyclic correlation matrix* of $\mathbf{x}(t)$ is used to find the correct direction, θ_s [1]. The cyclic correlation matrix has the following definition:

$$R_{xx}^{\alpha}(\tau) =< \mathbf{x}(t+\tau/2)\mathbf{x}^H(t-\tau/2)\cdot\exp(-j2\pi\alpha t)\} >_{\infty} . \quad (6)$$

If $R_{xx}^{\alpha}(\tau)$ exists and is nonzero for some τ and some nonzero α, then $\mathbf{x}(t)$ exhibits second order periodicity. Notice that $R_{xx}^{\alpha}(\tau)$ will be different for different τ. In [1], no value of τ is specified for use in cyclic-MUSIC. It is natural, then, to seek ways of using the additional dimension to improve the performance by including multiple lags of $R_{xx}^{\alpha}(\tau)$.

2.3 Signal Extraction by SCORE

Another array processing problem is from a signal extraction point of view. In this approach, the receiver can use a priori information concerning the statistics of the SOI to separate it from noise and interference. In the case of SCORE, it is assumed that the receiver knows that the SOI has non-zero cyclostationary content for some τ at a known cyclic frequency, α.

The extraction of the SOI is accomplished by constructing a suitable cost function on $y(t)$ and optimizing with respect to the weights. In the least-squares SCORE algorithm, the following cost is used [2]

$$F_{sc}(\mathbf{w};\mathbf{c}) =< |y(t) - r(t,\tau)|^2 >_N \quad (7)$$

where $r(t,\tau) = \mathbf{c}^H\mathbf{x}(t-\tau)e^{j2\pi\alpha t}$ and N is the data length. The solution to (7) can be thought of as maximizing the correlation between $\mathbf{x}(t)$ and a time- and frequency-shifted version of itself. This correlation can only exist for a nonzero α if $\mathbf{x}(t)$ exhibits second-order periodicity.

It is shown in [2] that the solution which minimizes (7) (assuming infinite data is available) is

$$\mathbf{w}_{opt} = R_{xx}^{-1}R_{xx}^{\alpha}(\tau)\mathbf{c}e^{-j\pi\alpha\tau} \quad (8)$$

where $\mathbf{c}$ is an arbitrary control vector as long as $\mathbf{c}^H\mathbf{a}(\theta_s) \neq 0$. It is shown in [2] that (8) maximizes the receiver's output SNR. However, the true solution may take long data lengths to compute accurately. Therefore, it is highly desirable to look for ways to improve SCORE's convergence properties. Intelligent use of the additional information given at different τ gives us a more robust SCORE algorithm.

3 Robustified Algorithms

3.1 Multiple-τ Subspace DOA Estimation

Consider an environment consisting of a single plane wave signal, $s(t)$, embedded in noise. Our signal model is

$$\mathbf{x}(t) = \mathbf{a}(\theta_s)s(t) + \mathbf{n}(t). \quad (9)$$

The goal, then, is to estimate θ_s from $\mathbf{x}(t)$.

By obtaining L cyclic correlation matrices corresponding to L different τ, L estimates of the null space can be obtained. The L null space estimates can be combined to produce a more accurate estimate of the true space.

It is proposed here to first find the null space of each cyclic correlation matrix available, according to how many τ are used. Second, stack these matrices to form a new matrix, V_{nn}:

$$V_{nn} = [\ V_n(\tau_1) \quad V_n(\tau_2) \quad \cdots \quad V_n(\tau_L)\ ]. \quad (10)$$

Here, $V_n(\tau_i)$ represents the null space estimate based on the cyclic correlation matrix corresponding to $\tau = \tau_i$. The new null space is formed by computing the left singular vectors of V_{nn} corresponding to the largest singular values. These vectors will be the basis of the new estimated null space. If L different values of τ are used, then the columns of the new matrix, say U_L, will be the singular vectors corresponding to the $n-M$ largest singular values of V_{nn}, where M is the number of signals. We can now estimate θ_s using

$$\hat{\theta}_s = \max_{\theta} \frac{1}{||U_L^H\mathbf{a}(\theta)||}. \quad (11)$$

Similarly, we may estimate the DOA from the signal subspace, which is found by seeking the M dominant left singular vectors of

$$\mathbf{V}_{ss} = [\ \mathbf{V}_s(\tau_1) \quad \mathbf{V}_s(\tau_2) \quad \cdots \quad \mathbf{V}_s(\tau_L)\ ]. \quad (12)$$

3.2 Multiple-τ SCORE

The reason for the slow convergence of SCORE is the time necessary to compute, by averaging incoming data, the matrices R_{xx} and $R_{xx}^{\alpha}(\tau)$. The accuracy of both improve as more data is averaged. The more accurate the calculation of those two matrices are, the closer the weight vector given by SCORE will be to the optimum SINR solution, $\mathbf{w}_{opt}$. The calculation of R_{xx}^{-1} itself cannot be accelerated. $R_{xx}^{\alpha}(\tau)$ can be computed with greater precision, however, if information taken from different values of τ are combined.

3.2.1 Pure Average Multiple-τ SCORE

One approach is to alter the existing SCORE cost function. We will show one obvious way of doing so with the least squares variety of SCORE.

The new cost that is proposed is

$$F_{m-sc}(\mathbf{w};\mathbf{c}) =< \sum_{\tau} |y(t) - r(t,\tau)|^2 >_N . \quad (13)$$

The optimization of F_{m-sc} will take advantage of the spectral correlation property of $\mathbf{x}(t)$ at various lags. Because the new cost is a quadratic function of $\mathbf{w}$, the global minimum can be found by taking the gradient of

$$F_{m-sc}(\mathbf{w};\mathbf{c}) = \mathbf{w}^H \hat{R}_{xx}\mathbf{w} + \mathbf{c}^H \hat{R}_{xx}\mathbf{c} \quad (14)$$
$$-\sum_{\tau}[\mathbf{c}^H \hat{R}_{xx}^{\alpha^*}(\tau)\mathbf{w}e^{-j\pi\alpha\tau} + \mathbf{w}^H \hat{R}_{xx}^{\alpha}(\tau)\mathbf{c}e^{-j\pi\alpha\tau}].$$

In the above expression, the $\hat{R}$ designation indicates an estimate from finite data. Taking the gradient with respect to $\mathbf{w}$, we get

$$\hat{R}_{xx}\mathbf{w} - \sum \hat{R}_{xx}^{\alpha}(\tau)\mathbf{c}e^{-j\pi\alpha\tau} = 0. \qquad (15)$$

This implies $\hat{\mathbf{w}}_{m-\tau} = \hat{R}_{xx}^{-1}[\sum_{\tau} \hat{R}_{xx}^{\alpha}(\tau)e^{-j\pi\alpha\tau}]\mathbf{c}$ where $\hat{\mathbf{w}}_{m-\tau}$ denotes the estimate of the optimum weight vector in the sense defined above. Notice that this is simply the sum of SCORE solutions taken at various τ.

We expect this method to be suboptimal. Consider that for some values of τ, the cyclic content of the signals may be zero. Therefore, data taken from lags which knowingly contain no signal information shouldn't contribute to the overall estimate. It stands to reason, then, that there exists a better method of combining the information. This reasoning prompted the following parameter estimation approach.

3.2.2 Minimum Variance SCORE Estimation

Recall that, assuming $s(t)$ is the only signal in the environment with cyclic content at frequency α, then

$$R_{xx}^{\alpha}(\tau) = r_{ss}^{\alpha}(\tau)\mathbf{a}(\theta_s)\mathbf{a}^H(\theta_s). \qquad (16)$$

Substituting (16) into (8), we have

$$\mathbf{w}_{opt} = r_{ss}^{\alpha}(\tau)R_{xx}^{-1}\mathbf{a}(\theta_s)\mathbf{a}^H(\theta_s)\mathbf{c}e^{-j\pi\alpha\tau}. \qquad (17)$$

Hence, $\mathbf{w}_{opt} = \gamma R_{xx}^{-1}\mathbf{a}(\theta_s)$ where γ is a complex scalar. It stands to reason that we can obtain a good estimate of $\mathbf{w}_{opt}$ by having a good estimate of $\mathbf{a}(\theta_s)$. This was the approach taken in [3], where the Cyclic Adaptive Beamforming (CAB) algorithm was proposed. In CAB, $\mathbf{a}(\theta_s)$ is obtained from measuring $R_{xx}^{\alpha}(\tau)$ and computing its left singular vector corresponding to its largest singular value. Asymptotically, this vector will be proportional to $\mathbf{a}(\theta_s)$. The constant of proportionality doesn't affect the output SINR of the array. However, the CAB algorithm does not give the optimum solution when the noise is not white.

We will approach this problem by determining the set of parameters $\{\beta(\tau)\}$ which give the smallest variance for a linear unbiased estimate of $\mathbf{a}(\theta_s)$ using $\{\hat{R}_{xx}^{\alpha}(\tau)\mathbf{c}\}_{\tau\in S}$ as the data set. S is the indexed set of all lags, τ, to be used for this purpose. Our estimate of $\mathbf{a}(\theta_s)$ will be

$$\hat{\mathbf{a}}(\theta_s) = \sum_{\tau\in S}\beta(\tau)\hat{R}_{xx}^{\alpha}(\tau)\mathbf{c}. \qquad (18)$$

To meet the unbiased requirement, $\{\beta(\tau)\}$ must satisfy

$$E\{\sum_{\tau\in S}\beta(\tau)\hat{R}_{xx}^{\alpha}(\tau)\mathbf{c}\} = \mathbf{a}(\theta_s) \qquad (19)$$

and it must minimize the variance of $\hat{\mathbf{a}}(\theta_s)$.

From (16) and (18), we see that

$$E\{\hat{a}(\theta_s)\} = E\{\sum_{\tau}\beta(\tau)[r_{ss}^{\alpha}(\tau)\mathbf{a}(\theta_s)\mathbf{a}^H(\theta_s) + \Delta(\tau)]\mathbf{c}\}$$
$$= \mathbf{a}(\theta_s)\mathbf{a}^H(\theta_s)\mathbf{c}\sum_{\tau}\beta(\tau)r_{ss}^{\alpha}(\tau) \qquad (20)$$

where $\hat{R}_{xx}^{\alpha}(\tau) = R_{xx}^{\alpha}(\tau) + \Delta(\tau)$ and $\Delta(\tau)$ is, by assumption, a zero mean, uncorrelated random perturbation matrix. That is, the elements of $\Delta(\tau)$ are uncorrelated with each other and uncorrelated over τ. (In reality, this assumption is somewhat imprecise, but simulation results indicate it has merit.) If the estimate is to be unbiased, then

$$|\mathbf{a}^H(\theta_s)\mathbf{c}\sum_{\tau}\beta(\tau)r_{ss}^{\alpha}(\tau)|^2 = 1. \qquad (21)$$

It should be noted, however, that the above quantity can be any nonzero number and still result in an estimate which is unbiased to a *scalar* multiple of $\mathbf{a}(\theta_s)$. This is important because scalar multiples of $R_{xx}^{-1}\mathbf{a}(\theta_s)$ are all optimum solutions to the SCORE algorithm. Proceeding, we have

$$var(\hat{\mathbf{a}}(\theta_s)) = \sum_{\tau_1}\sum_{\tau_2}\beta^*(\tau_1)\beta(\tau_2)\mathbf{c}^H E\{\Delta^H(\tau_1)\Delta(\tau_2)\}\mathbf{c}.$$

Letting the variance of the elements of $\Delta(\tau)$ be σ_{Δ}^2,

$$E\{\Delta^H(\tau_1)\Delta(\tau_2)\} = n\sigma_{\Delta}^2 I_{n\times n}\delta(\tau_1-\tau_2) \qquad (22)$$

Substituting (22) into the variance expression,

$$var(\hat{\mathbf{a}}(\theta_s)) = n\sigma_{\Delta}^2\|\mathbf{c}\|^2\sum_{\tau}|\beta(\tau)|^2. \qquad (23)$$

The optimization problem is to minimize (23) subject to (21). Taking an equivalent interpretation, we can maximize

$$\max_{\{\beta(\tau)\}} F_{\beta} = \frac{|\mathbf{a}^H(\theta_s)\mathbf{c}|^2|\sum_{\tau}\beta(\tau)r_{ss}^{\alpha}(\tau)|^2}{n\sigma_{\Delta}^2\|\mathbf{c}\|^2\sum_{\tau}|\beta(\tau)|^2}. \qquad (24)$$

We can use the Cauchy-Schwartz inequality on the numerator of F_β to write

$$F_\beta \leq \frac{|\mathbf{a}^H(\theta_s)\mathbf{c}|^2 \sum_{\tau_1} |\beta(\tau_1)|^2 \sum_{\tau_2} |r_{ss}^{\alpha}(\tau_2)|^2}{n\sigma_\Delta^2 \|\mathbf{c}\|^2 \sum_\tau |\beta(\tau)|^2} \quad (25)$$

where equality holds if and only if

$$\beta(\tau) = \gamma(r_{ss}^{\alpha}(\tau))^* \quad (26)$$

where γ is an arbitrary nonzero real scaling factor.

3.2.3 Efficiency of MV-SCORE

If we further assume that the elements of the matrix $\Delta(\tau)$ are i.i.d. gaussian, then the data becomes

$$\{\hat{R}_{xx}^{\alpha}(\tau)\mathbf{c}\}_{\tau\in S} = \{\lambda \mathbf{a}(\theta_s) + \mathbf{e}\}_{\tau \in S} \quad (27)$$

where $\mathbf{e}$ is a gaussian random vector with i.i.d. elements. Therefore, the minimum variance estimate would be efficient. For real data, the central limit theorem tells us that the minimum variance estimate will be asymptotically efficient.

4 Simulations

A. Multiple-τ Subspace Simulations:
In this simulation the signal environment consists of two PAM signals (of rate $1/T$) embedded in white, additive sensor noise arriving at angles of 40 and 50 degrees. The signals have cyclic content at the same frequency–both being over-sampled 5 times per symbol. The receiver platform is a linear array of 4 monopole antennas. The data length is 800 symbols and the input SNR is 0 dB.

Figure 1 shows the curves defined by (11) from sweeps of θ over 10 trials. Four different lags, τ, ($\tau = \{0, 1, 2, 3\}$) were used in the multiple-τ algorithm. The lag unit is $T/5$. The multiple-τ cyclic MUSIC method (top) produces estimates that have much smaller variance to those given by the single-τ ($\tau = 0$) algorithm (bottom).

B. SCORE Simulations:
In Figure 2, there is a single signal in the environment arriving at 40 degrees embedded in white noise. The input SNR is 10 dB. The graph represents the output SNR of the SCORE algorithm corresponding to different τ. An average of 10 trials is shown. The performance of SCORE is much different depending on which τ is chosen. The pure average and minimum variance estimators used information from $\tau = \{0, 2, 4, 6, 8\}$. The simulation indicates that the minimum variance method outperforms a pure average as well as the single-τ methods in both long and short terms.

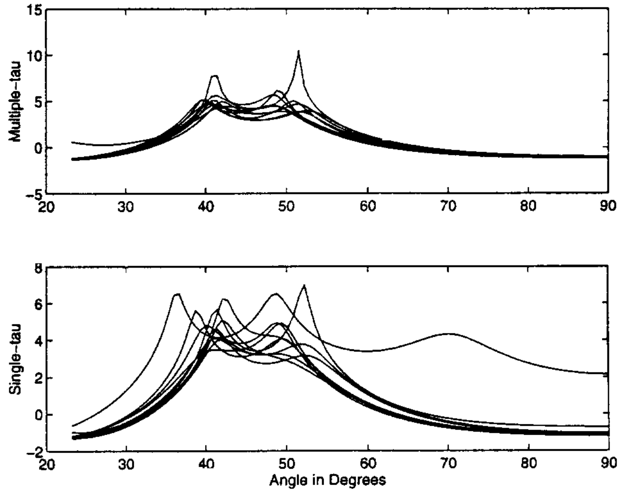

Figure 1: Cyclic MUSIC comparison.

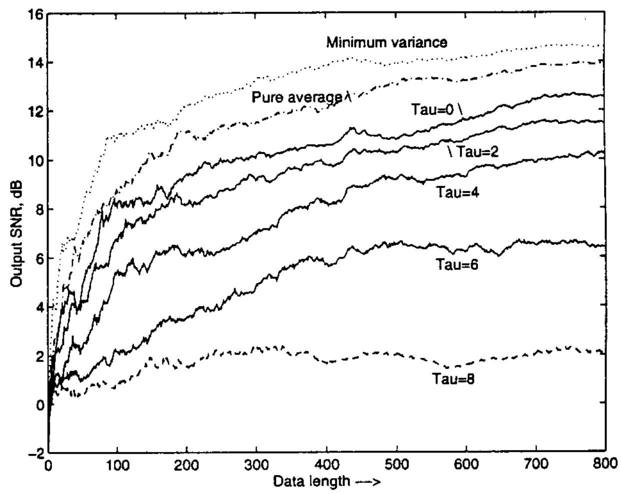

Figure 2: SCORE comparisons.

References

[1] W. A. Gardner, "Cyclic MUSIC Algorithms for Signal-Selective Direction Estimation", *Proc. ICASSP*, pp. 2278-2281, 1989.

[2] B. G. Agee, S. V. Schell, W. A. Gardner, "Spectral Self-Coherence Restoral: A New Approach to Blind Adaptive Signal Extraction Using Antenna Arrays", *Proc. IEEE*, pp. 753-767, April 1990.

[3] Q. Wu, K. M. Wong, R. Ho, "A Fast Algorithm for Adaptive Beamforming of Cyclic Signals," *ICWC*, 1993.