The maximum entropy method
and its application
to clutter cancellation

EGIDIO D'ADDIO - ALFONSO FARINA - FLAVIO A. STUDER
Radar Dept. - System Analysis

SUMMARY:

Two high resolution techniques for spectral estimation of random processes are described. The estimate is obtained from samples of the autocorrelation function or directly from samples of the process. These techniques are based on the Maximum Entropy Method (MEM).

After a review of the main points of this method, an application to radar system is shown in detail. First, the estimation of clutter spectrum is considered; then, this estimate is exploited to shape a filter for clutter cancellation and target echo enhancement.

The processing algorithm is an adaptive one, and its performances are evaluated, by means of computer simulation, in term of Improvement Factor and speed of adaptation.

1. Introduction

The choice of the architecture and the parameters of a signal processor are of particular interest in the design of a radar system that has to operate in a particularly severe natural or artificial interference. An outstanding example of this are radars operating in an environment in which a combination of diverse types of clutter are present (ground, sea, rain) whose time evolution is not known. As a result, the use of an adaptive processor is necessary, one that evaluates the statistical characteristics of the disturbance and consequently performs a filtering action that attenuates the clutter and contemporaneously enhances the useful signal.

A theory to design a filter structure that maximizes the output signal/noise ratio exists in literature [1 •• 6]. It is important to emphasize that such a theory is absolutely general and not restricted to a particular type of radar. The optimum processor is a transversal filter (FIR: Finite Impulse Response) that linearly combines N samples of the received signal. The filter output envelope is extracted and compared with a suitable threshold.

A functional diagram of the optimum processor is shown in figure 1. The linear combination coefficients (weights) \( W_i \) (i = 1, 2, ..., N) are provided in vectorial notation by the following expression:

\[
\mathbf{W} = \mu \mathbf{M}^{-1} \mathbf{S^*}
\]  

where \( \mu \) is an arbitrary constant, \( \mathbf{M} \) is the disturbance covariance matrix and \( \mathbf{S} \) is the vector that represents the expected useful signal samples.

The filter output is equal to \( \mathbf{Z}^T \mathbf{M}^{-1} \mathbf{S^*} \) in which \( \mathbf{Z} \) indicates the vector containing the N input signal samples.

This output can be interpreted as the result of two cascaded operations: the multiplication of \( \mathbf{Z}^T \) by \( \mathbf{M}^{-1} \) corresponds to a transformation of the disturbance correlated component (clutter) into white noise; the result of this operation is then multiplied by \( \mathbf{S^*} \) to enhance the useful signal (matched filter), if present.

An adaptive realization of the filter in fig. 1 is based on the dynamic updating of the weights \( W_i \) as a function of the disturbance parameters, synthesized in matrix \( \mathbf{M} \). The elements of this matrix are obtained from the autocorrelation function of the disturbing process whose Fourier Transform is the power spectral density. Therefore, the link existing between the optimum adaptive filtering and the disturbance spectral estimation is clear: from the spectral estimation the autocorrelation function can be obtained and then weights for the optimum filtering can be derived.

Independently of such an application, the spectral estimation problem has an inherent importance in signal
processing in many fields: communications, astronomy, geophysics. Therefore it has been largely analyzed in the last few years.

The first method, widely used to estimate a signal spectrum from a finite number of its samples, is the method of periodograms [7]; from a statistical viewpoint it produces unsatisfactory results because the obtained estimate is affected by errors comparable with the estimated value. The fundamental limit of this method is that it is based on a time average carried out on a limited number of data, whose result differs notably from the statistical average. Furthermore, as will be better shown in paragraph 2, outside of the time window of observation such a method induces to make arbitrary assumptions on the process itself.

More recently the maximum entropy method (MEM) has been successfully applied to the spectral estimation of stochastic processes.

This algorithm, based on linear prediction theory, permits the extension of the number of available data by extrapolating the process beyond the observation interval.

In this paper, the MEM method is examined in detail as a high resolution estimator of a clutter spectrum and as a means of implementing eq. LA brief review of the entropy concept as used in information theory is made in paragraph 2. In the following sections, this concept is applied to spectral estimation starting from a limited number of samples of the autocorrelation function, or directly from samples of the process itself. Finally, a possible use of the maximum entropy method for the implementation of the optimum filter is analyzed and its performance are evaluated.

2. The entropy concept in information theory

Consider a system having uncertain characteristics whose behavior can be described by means of the occurrence of an event $E_j$ amongst $N$ possibilities, each one having an a priori probability $P_j$.

To the occurrence of an event $E_k$ is associated a quantity of information $I_k$ that depends on $P_k$ according to the relation:

$$I_k = - \log_2 P_k$$

(2)

With such a definition the occurrence of independent events results in information that is the sum of those of the single events; besides, the least probable event gives the maximum information. If the system is observed for a time $T$, the value of the total information obtained in the whole period is:

$$I = - \sum_{i=1}^{N} P_i \cdot T \cdot \log_2 P_i$$

(3)

The system entropy is defined as the average information per unit time, as follows:

$$H = \frac{I}{T} = - \sum_{i=1}^{N} P_i \cdot \log_2 P_i$$

(4)

It is therefore a measure of the uncertainty described by the set of the a priori probabilities.

As a consequence, the entropy is zero if a certain $P_i$ is equal to 1 and all the others are zero. On the other hand, if all the $P_i$ are equal, the entropy is maximum.

With these assumptions, consider now a situation in which a random variable $X$ can assume the values $X_1, X_2, X_3, ..., X_N$. Besides, assume that a certain quantity of information on $X$ is given, for example the knowledge of the average values: $F_1, F_2, F_3, ..., F_m$ of a certain number of functions $f_1(X), f_2(X), ..., f_m(X)$ of the variable $X$.

It is interesting to choose in the best way (in a completely objective manner) the a priori probability distribution

$$P_i = \text{Prob} \{X=X_i\} \quad i = 1,2,...,N$$

utilizing exclusively the information that are given. A way to assign the distribution is based on the concept of entropy just introduced. In fact, if we find the $P_i$, $i = 1,2,...,n$ that maximize the system entropy $H$ defined by (4), subject to the conditions:

$$\sum_{i=1}^{N} P_i = 1$$

(5)

$$\sum_{k=1}^{m} P_i f_k(X_i) = F_k \quad k = 1,2,...,m$$

we obtain an a priori distribution to which corresponds the maximum possible entropy, subject to the constraints (5).

Therefore we can conclude that:

"The assignment of the a priori probability that is in agreement with the problem data but is the most possibly vague concerning the unknown information is the one to which corresponds the maximum entropy".

The concept just stated is applied to the signal spectrum estimation in the following sections.

3. Maximum entropy spectral estimation method applied to a finite number of values of the autocorrelation function.

The traditional spectral estimation methods are based on the calculation of samples of the autocorrelation function that when transformed gives an estimate of the power spectrum. Nevertheless, as the autocorrelation is only known in a limited interval, it is generally assumed null or periodic outside.

In addition, to reduce the sidelobes of the spectral estimation, the data are weighted so that the autocorrelation goes to zero with continuity. This approach does not make an objective use of the a priori information in that it alters the data and arbitrarily extend them beyond the observation interval. Such methods have a resolution that is limited by the number of samples that are available, and because this in general is limited (especially in radar application), the resolution is poor.
In order to obtain a method with a greater resolution, in 1967 J.P. Burg proposed the use of MEM. The situation under consideration lends itself well to the application of maximum entropy, in fact the a priori available information are \((N + 1)\) values of the autocorrelation function \(R(\tau)\) (eventually obtainable from the signal samples) of the process whose spectrum is to be estimated. The available data concern the II order moment; it is possible to show that amongst all the processes with the same I and II order moments the Gaussian process has the maximum entropy.

Consequently, it is assumed that the original process has a Gaussian probability density to which corresponds an entropy \([8,9]\):

\[
H = \frac{1}{4f_N} \int_{-f_N}^{f_N} \log_2 S(f) \, df
\]

where \(f_N = \frac{1}{2\Delta t}\) with \(\Delta t\) sample period and \(S(f)\) spectrum that is to be estimated. The problem can therefore be formulated in the following manner:

- Calculate \(S(f)\) that maximizes (6) with the following constraints:

\[
R(n) = \int_{-f_N}^{f_N} S(f) \exp(-2\pi jfn\Delta t) \, df
\]

\(n = -N, \ldots, N\)

Applying the Lagrange multipliers method to solve this problem, the following fundamental expression for the estimated spectrum is obtained

\[
S_{MEM}(f) = \frac{\Delta t \cdot P}{|1 + \sum_{k=1}^N W_k \exp(-j2\pi kf\Delta t)|^2}
\]

where \(P\) and the coefficients \(W_k\) are calculated by solving the following matrix systems

\[
\begin{bmatrix}
R(0) & R(1) & \ldots & R(N) \\
R(1) & R(0) & \ldots & R(N-1) \\
R(N) & R(N-1) & R(0)
\end{bmatrix}
\begin{bmatrix}
1 \\
W_1
\end{bmatrix}
= \begin{bmatrix}
P \\
0
\end{bmatrix}
\]

(9) is rewritten

\[
\begin{bmatrix}
R(0) & R(1) & \ldots & R(N-1) \\
R(1) & R(0) & \ldots & R(N-2) \\
R(N-1) & R(N-2) & \ldots & R(0)
\end{bmatrix}
\begin{bmatrix}
1 \\
W_1
\end{bmatrix}
= \begin{bmatrix}
P \\
0
\end{bmatrix}
\]

that generates the following system of \(N + 1\) equations

\[
\begin{cases}
P = R(0) + R^T W \\
W = -M_R^{-1} R
\end{cases}
\]

At this point one becomes aware that the spectral estimation obtained by the maximum entropy method corresponds to make a fitting of the autocorrelation function by means of an order \(N\) autoregressive process in the \((N + 1)\) time instants in which it is known and to extrapolate the process when it is unknown using the same autoregressive model.

In order to show what has been said it is necessary to briefly recall linear prediction (cfr. [8,9]) and autoregressive processes. Given a process \(X(t)\) it is required to estimate the \((N + 1)\)th sample of the process by a linear combination of \(N\) previous known samples (forward predictor filter) (*):

\[
\hat{X}_{N}(n+1) = \hat{W}_1 X(n) + \hat{W}_2 X(n-1) + \ldots + \hat{W}_N X(n-N)
\]

\[
X(n-N+1) = \sum_{k=1}^N \hat{W}_k X(n-k+1)
\]

where with \(\hat{X}_{N}(n)\) is indicated the estimate of \(X(n)\) using \(N\) samples. The value of the estimation error is:

\[
\hat{e}_N(n) = X(n) - \hat{X}_{N}(n)
\]

An order \(N\) prediction filter scheme is shown in figure 2. The calculation of the weights \(\hat{W}_k\) is made by imposing that the mean square error \(e^2\) be a minimum. It can be shown that (Ref. [2,8]) the weights \(\hat{W}_k\) are obtained by solving the matrix relation (9) or the equivalent (14) where \(R(0), R(1), \ldots, R(N)\) represent the assumed known \((N + 1)\) values of the process autocorrelation function and \(P\) is the

\[
R^T = [R(1), R(2), \ldots, R(N)]
\]

\[
W^T = [W_1, W_2, \ldots, W_N]
\]

\[
M_R = \begin{bmatrix}
R(0) & R(1) & \ldots & R(N-1) \\
R(1) & R(0) & \ldots & R(N-2) \\
R(N-1) & R(N-2) & \ldots & R(0)
\end{bmatrix}
\]

\[
\begin{bmatrix}
1 \\
W_1
\end{bmatrix}
= \begin{bmatrix}
P \\
0
\end{bmatrix}
\]

\[
\hat{X}_{N}(n+1) = \hat{W}_1 X(n) + \hat{W}_2 X(n-1) + \ldots + \hat{W}_N X(n-N)
\]

\[
X(n-N+1) = \sum_{k=1}^N \hat{W}_k X(n-k+1)
\]

\[
\hat{e}_N(n) = X(n) - \hat{X}_{N}(n)
\]

(*): It is possible to define a "backward predictor filter" estimating \(X(n)\) on the basis of \(N\) successive samples of the process. Such a filter has a transfer function: \(H(z) = H(1/z^*).\)
mean square value of the prediction error e(n). Consider the error \( \hat{e}(n) \) as the output of the said «forward error filter», the transfer function from \( X(n) \) to \( \hat{e}(n) \) is:

\[
\hat{H}(z) = 1 - \sum_{k=1}^{N} \hat{W}_k Z^{-k}
\]  

(17)

Furthermore it is shown that the sequence \( \hat{e}(n) \) is white, therefore the filter \( H(z) \) is «whitening» for the input process \( X(n) \). A process is said to be autoregressive, with order \( N \), if each sample is obtained by a linear combination of \( N \) previous samples and a white Gaussian noise with zero mean and variance \( \sigma^2 \), that is

\[
X(n) = \sum_{k=1}^{M} C_M(k) X(n-k) + e(n)
\]  

(18)

and corresponds to a filter with a transfer function from input \( \hat{e}(n) \) and output \( X(n) \)

\[
T(z) = \frac{1}{1 - \sum_{k=1}^{N} c_k Z^{-k}}
\]  

(19)

from a comparison with (17) is found that

\[
T(z) = 1/\hat{H}(z)
\]  

(20)

The power spectrum \( S(\omega) \) of the autoregressive process \( X(n) \) is

\[
S(\omega) = \frac{\sigma^2}{|\hat{H}(z)|^2} = \frac{\sigma^2}{1 - \sum_{k=1}^{N} C_k \exp(-jk\omega\Delta t)}
\]  

(21)

where \( \Delta t \) is the sampling interval.

As the following relation holds \( \sigma^2 = \Delta t P \) where \( P \) is the prediction filter mean square error and \( f_N = \frac{1}{2\Delta t} \) is the process sampling frequency, if \( C_k = W_k \), expression (21) is perfectly similar to (8). It is demonstrated that the MEM spectral estimation corresponds to generating an autoregressive process that performs the data fitting and extrapolation.

Let us now evaluate the performance of the MEM algorithm as a spectrum estimator: if the algorithm provides sufficient accuracy and resolution as a spectrum estimator, it also allows an appreciable clutter cancellation. This evaluation is carried out by means of simulation; \( N \) samples of an autocorrelation function, relative to thermal noise and to one or two clutters with Gaussian spectrum, are generated and the overall spectrum is estimated applying eqs. (14) and (21).
In figures 4a - 4b are shown the results of the method applied to the spectrum of one or two Gaussian clutters in white noise whose average power is 30 dB lower than the clutter. In figure 4a are shown 4 estimates for a single clutter, with correlation coefficient equal to 0.9, by means of a number of known values of the process autocorrelation function equal to 2, 3, 4, and 5: this figure shows that 4 samples are sufficient to obtain a satisfactory estimate.

In figure 4b the estimate for the case of two clutters of equal average power, positioned around zero frequency and at half of the PRF, respectively, is shown.

In this case, the number of necessary samples is greater, due to the very stringent condition of two clutters of same power. In particular, with only two samples the information is so poor that the estimate yields a white spectrum that tends to average out the two real spectra. Nevertheless, with nearly 10 samples it is possible to obtain good results.

4. The maximum entropy method applied to signal samples

A more significant case for the spectral estimate is when N process samples are given from which it is required to obtain the MEM power spectrum estimate. The problem is to evaluate the coefficients $C_M(k)$ and the order $M$ of an autoregressive filter of the following type (cfr. eq. (18)):

$$X(n) = \sum_{k=1}^{M} C_M(k) X(n-k) + e(n)$$

from the process known samples, and can be solved by using the Burg recursive algorithm [7].

In order to explain how the recursive algorithm works consider that the autoregressive process (22) can be seen as the output of a «forward» prediction filter with coefficients $[1, -C_M(1), -C_M(2), \ldots, -C_M(M)]$ that moves on the data string from left to right for increasing time. Obviously one can also refer to a «backward» filter with coefficients $[1, -C^*_M(1), -C^*_M(2), \ldots, -C^*_M(M)]$ that moves in decreasing time on the data (cfr. figure 5).

The simultaneous use of these two filters permits the complete use of all the data available. The coefficients
\( C_M(i) \) are calculated by minimizing the power of the total error made by the forward and backward filters to predict the available data. Let \( f_M(k) \) and \( b_M(k) \) be the prediction error of the forward and backward filter respectively, the total prediction error energy is:

\[
\hat{e}^2(M) = \sum_{k=1}^{NM} |f_M(k)|^2 + |b_M(k)|^2
\]

that is to be minimized with respect to coefficients \( C_M(k) \).

The algorithm is based on a set of recursive equations relating the prediction errors and the coefficients:

\[
C_M(k) = C_{M-1}(k) + C_M(M) C^*_{M-1}(M-k)
\]

\[
\hat{e}^2(M) = [1 - |C_M(M)|^2] \hat{e}^2(M-1)
\]

\[
f_M(k) = f_{M-1}(k) + C_M(M) b_{M-1}(k-1)
\]

\[
b_M(k) = b_{M-1}(k) + C^*_{M}(M) f_{M-1}(k)
\]

Therefore the only parameter to calculate is the coefficient \( C_M(M) \) having the following expression:

\[
C_M(M) = \frac{-2 \sum_{k=1}^{M} b_{M-1}(k) f_{M-1}(k+1)}{\sum_{k=1}^{M} [b_{M-1}(k)]^2 + [f_{M-1}(k+1)]^2} = \frac{2 \text{NUM}}{\text{DEN}}
\]

from which it results \( |C_M(M)| \leq 1 \) hence the error \( e(M) \) decreases as the prediction filter order increases and the algorithm is stable.

Now, it is necessary to specify how to solve the initialization problem and to assign the maximum order to the filter. The initialization is brought about in the following manner:

\[
\hat{e}^2(0) = \frac{1}{N} \sum_{k=1}^{N} |X(k)|^2
\]

\[
f_0(k) = b_0(k) = X(k) \quad k = 1, 2, \ldots, N
\]

The choice of the filter maximum order \( M \) is a critical question in that if it is too low a poor resolution is obtained; if vice versa it is too high, the errors made in the estimate of the coefficients are accumulated and can induce spurious peaks in the spectrum. An often used criterion for the selection of \( M \) is due to Akaike [11] and is known as the Final Predictor Error (FPE).
It can be shown that FPE relative to $N$ samples and to a filter of order $M$ is:

$$FPE(M) = \frac{N \times (M+1)}{N \times (M+1)^2} e(M)$$  \hspace{1cm} (29)$$

the choice of $M$ can be made by applying the Burg algorithm repeatedly to the same group of $N$ samples with increasing order up to $N/2$ (in order not to use overlapping data) and selecting the value that minimizes FPE ($M$). Recent experience in the application of the algorithm in the radar field (cfr. [12]) and simulation results show that $M = N/2$ is a convenient choice.

In conclusion, the MEM spectral estimator is:

$$S(\omega) = \frac{e(M) \Delta t}{|1+\sum_{k=1}^{N/2} C_k(M) \exp(-j\omega k \Delta t)|^2}$$ \hspace{1cm} (30)$$

In Fig. 6 Burg algorithm «flow chart» is shown.

Lastly, in order to improve the spectral estimation it is convenient to average the coefficient estimates using more independent sets of data, having the same statistical characteristics. In radar application this corresponds to consider some range bins around the one to be processed, as illustrated in fig. 7, where each $i$-th set of independent samples is $X_{1i}, X_{2i},....X_{Ni}, (i=1,2,...., NS)$.

It is necessary that the number of cells $NS$ be small enough to guarantee that the clutter remains stationary over the pertaining range interval.

In practice, the average is performed only on the coefficient $C_M(M)$ according to the expression

$$C_M(M) = \frac{-2 \sum_{k=1}^{MS} NUM(k)}{\sum_{k=1}^{M} DEN(k)}$$  \hspace{1cm} (31)$$

Also in this case, the Burg algorithm has been evaluated by simulation.

The result of a spectral estimation for one or two Gaussian clutters imbedded in thermal noise are shown in figure 8a and 8b. In the same figure the exact spectra are shown in dotted lines. When one clutter is present it is sufficient to process ten samples by means of a prediction filter of third order. In the case of two clutters and prediction filter of the tenth order are required in the particularly severe case of two clutters having equal power.

Finally, it can be seen that it is sufficient to average the estimate on ten sets of independent samples to obtain sufficiently accurate results.

5. Application of MEM to optimum clutter rejection

In this section the application of the Burg algorithm to clutter cancellation is described. The processing is carried out by using eqn. (1) without the inversion of the covariance matrix $M$ to obtain a noteable reduction of hardware complexity. In fact it is possible to show that the inverse of the covariance matrix $M$ can be directly calculated from the coefficients $C_M(1), C_M(2),...C_M(M)$ of the spectral estimate prediction filter and from the prediction error (Ref. [12] to [15]). To demonstrate it, we refer to the following example: given $N = 2$ clutter samples, with an average power $\sigma$ and correlation coefficient $\rho$, the corresponding covariance matrix is:

$$M = \sigma^2 \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$$  \hspace{1cm} (32)$$

In this simple case the Burg's algorithm reduces to the calculation of $C_1$ minimizing $\varepsilon_1$. From (33) we obtain:

$$\varepsilon_1 = \sigma^2 [1 + C_1(1)-2Q C_1(1)]$$  \hspace{1cm} (34)$$

minimizing the expression (34) with respect to $C_1(1)$ yields:

$$C_1(1) = Q$$  \hspace{1cm} (35)$$

$$\varepsilon_1 = \sigma^2 (1-Q^2)$$

Let now define the following $(2 \times 2)$ matrices:

$$A = \begin{bmatrix} 1 & 0 \\ -C_1(1) & 1 \end{bmatrix}$$
$$D^{-1} = \begin{bmatrix} \varepsilon_1 & 0 \\ 0 & \sigma^2 \end{bmatrix} = \sigma^2 \begin{bmatrix} 1-Q^2 & 0 \\ 0 & 1 \end{bmatrix}$$  \hspace{1cm} (36)$$

$$A^T D^{-1} A = \frac{1}{\sigma^2(1-Q^2)} \begin{bmatrix} 1 & -Q \\ -Q & 1 \end{bmatrix}$$  \hspace{1cm} (37)$$

Where $A$ is the filter coefficients matrix, $D$ is the prediction error matrix.

It can be verified that in general the following decomposition relation hold for $M^+$:

$$M^+ = A^T D^{-1} A = \frac{1}{\sigma^2(1-Q^2)} \begin{bmatrix} 1 & -Q \\ -Q & 1 \end{bmatrix}$$  \hspace{1cm} (38)$$

Therefore the clutter cancellation filter can be synthesized from the matrices $A$ and $D$. These generally have expression more complex than (36) and (37).

A simulation has been made, to evaluate the improvement factor obtainable for one or two clutters with Gaussian spectra, and for targets signal with a given doppler frequency.

The results of the simulation, shown in Table 1, are relative to a clutter with 30 dB power to noise ratio, centered on a zero frequency and with correlation coefficients $\rho = 0.9$ and 0.99.
Fig. 8a - MEM estimate of a Gaussian spectrum with $\rho = 0.9$ (dotted curve) with the Burg method: 10 samples, prediction order M, CNR = 30dB.

Fig. 8b - Analogous estimate to that of figure 8a with an average on 10 trials.

Fig. 8c - MEM spectrum estimate of two equal power Gaussian clutters utilizing the Burg method: 10 samples, prediction order M, CNR = 30dB.

Fig. 8d - Analogous estimate to that shown in figure 8c with an average on ten trials.
The target has a doppler frequency of 0.5 PRF c 0.25 PRF where PRF is the Pulse Repetition Frequency.

In Table 1 the obtained Improvement Factor is compared with the optimum one, for several values of the number of processed samples and different correlation coefficients. In all cases, the loss is within a few dBs. However, from these results, it is not possible to fit a relation between the loss and the varying parameters. This is due to the fact that, for sake of simplicity, the degree of the prediction filters is not optimized, but is set at N/2 in all conditions. As a general conclusions, the obtained performance are reasonably close to the optimum, provided that an average is made on about 10 range-cells.

In Table 2 are shown the results relative to two clutters, having different central frequencies, f1 and f2, correlation coefficients and power; also in this case, for any values of target doppler frequency, the Improvement Factor is close to the optimum, when an avarage is made on at least 10 cells.

6. Conclusions

The application of the maximum entropy principle to clutter spectrum estimation and to the design of the optimum filter for its cancellation has been shown. The illustrated algorithms performance have been evaluated by means of simulation and calculation programs.

A deeper study in the application of this method to radar signal processing is encouraged by the obtained results. In particular it is necessary to find an efficient algorithm application scheme and to evaluate its sensitivity to the accuracy of the data representation.

Finally, a performance comparison between the optimum filter, implemented by means of MEM, and more conventional processors (MTI, adaptive MTI, FFT), is going to be made in a next paper to be issued. The aim is to recognize in what cases the fully adaptive optimum filter, based on MEM principle, is a cost/effective solution.

| TABLE 1 |
| The optimum filter performance obtained with MEM for a Gaussian clutter with CNR = 40dB. Thr average is carried out on ten trials. |

<table>
<thead>
<tr>
<th>( \rho )</th>
<th>( f_d/PRF )</th>
<th>( N^* ) PULSE</th>
<th>OPTIMUM IMP. (dB)</th>
<th>IMP. OBTAINED (dB)</th>
<th>LOSS (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>.90</td>
<td>.500</td>
<td>4</td>
<td>29.86</td>
<td>25.97</td>
<td>.29</td>
</tr>
<tr>
<td>.90</td>
<td>.500</td>
<td>6</td>
<td>35.93</td>
<td>34.69</td>
<td>1.24</td>
</tr>
<tr>
<td>.90</td>
<td>.500</td>
<td>8</td>
<td>37.57</td>
<td>34.17</td>
<td>3.40</td>
</tr>
<tr>
<td>.90</td>
<td>.500</td>
<td>10</td>
<td>38.87</td>
<td>33.08</td>
<td>5.79</td>
</tr>
<tr>
<td>.90</td>
<td>.250</td>
<td>4</td>
<td>20.82</td>
<td>16.80</td>
<td>4.02</td>
</tr>
<tr>
<td>.90</td>
<td>.250</td>
<td>6</td>
<td>24.02</td>
<td>22.25</td>
<td>1.77</td>
</tr>
<tr>
<td>.90</td>
<td>.250</td>
<td>8</td>
<td>26.00</td>
<td>21.74</td>
<td>4.26</td>
</tr>
<tr>
<td>.90</td>
<td>.250</td>
<td>10</td>
<td>27.28</td>
<td>19.49</td>
<td>7.79</td>
</tr>
<tr>
<td>.95</td>
<td>.500</td>
<td>4</td>
<td>34.01</td>
<td>33.74</td>
<td>.28</td>
</tr>
<tr>
<td>.95</td>
<td>.500</td>
<td>6</td>
<td>36.44</td>
<td>36.28</td>
<td>.16</td>
</tr>
<tr>
<td>.95</td>
<td>.500</td>
<td>8</td>
<td>38.21</td>
<td>33.63</td>
<td>4.58</td>
</tr>
<tr>
<td>.95</td>
<td>.500</td>
<td>10</td>
<td>39.28</td>
<td>39.14</td>
<td>.14</td>
</tr>
<tr>
<td>.95</td>
<td>.250</td>
<td>4</td>
<td>25.77</td>
<td>22.89</td>
<td>2.87</td>
</tr>
<tr>
<td>.95</td>
<td>.250</td>
<td>6</td>
<td>30.55</td>
<td>27.81</td>
<td>2.73</td>
</tr>
<tr>
<td>.95</td>
<td>.250</td>
<td>8</td>
<td>33.68</td>
<td>24.67</td>
<td>9.01</td>
</tr>
<tr>
<td>.95</td>
<td>.250</td>
<td>10</td>
<td>35.91</td>
<td>30.87</td>
<td>5.03</td>
</tr>
<tr>
<td>.99</td>
<td>.500</td>
<td>4</td>
<td>35.08</td>
<td>34.56</td>
<td>.52</td>
</tr>
<tr>
<td>.99</td>
<td>.500</td>
<td>6</td>
<td>37.33</td>
<td>28.76</td>
<td>8.58</td>
</tr>
<tr>
<td>.99</td>
<td>.500</td>
<td>8</td>
<td>38.62</td>
<td>27.94</td>
<td>10.68</td>
</tr>
<tr>
<td>.99</td>
<td>.500</td>
<td>10</td>
<td>39.64</td>
<td>36.35</td>
<td>3.29</td>
</tr>
<tr>
<td>.99</td>
<td>.250</td>
<td>4</td>
<td>31.83</td>
<td>31.11</td>
<td>.72</td>
</tr>
<tr>
<td>.99</td>
<td>.250</td>
<td>6</td>
<td>36.09</td>
<td>30.41</td>
<td>5.68</td>
</tr>
<tr>
<td>.99</td>
<td>.250</td>
<td>8</td>
<td>38.21</td>
<td>31.00</td>
<td>7.20</td>
</tr>
<tr>
<td>.99</td>
<td>.250</td>
<td>10</td>
<td>39.31</td>
<td>33.85</td>
<td>5.46</td>
</tr>
</tbody>
</table>
TABLE 2
Optimum MEM filter performance for two Gaussian clutters, CNR1 = 40dB, CNR2 = 20db, average on ten trials.

<table>
<thead>
<tr>
<th>CNR2 (dB)</th>
<th>$\rho_1$</th>
<th>$\rho_2$</th>
<th>$f_s$/PRF</th>
<th>$f_d$/PRF</th>
<th>N* PULSE</th>
<th>OPTIMUM IMP. (dB)</th>
<th>IMP. OBTAINED (dB)</th>
<th>LOSS (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>.95</td>
<td>.90</td>
<td>.00</td>
<td>.50</td>
<td>.26</td>
<td>10</td>
<td>34.36</td>
<td>30.41</td>
</tr>
<tr>
<td>20</td>
<td>.95</td>
<td>.90</td>
<td>.00</td>
<td>.50</td>
<td>.26</td>
<td>14</td>
<td>39.25</td>
<td>33.20</td>
</tr>
<tr>
<td>20</td>
<td>.95</td>
<td>.90</td>
<td>.00</td>
<td>.50</td>
<td>.26</td>
<td>18</td>
<td>42.38</td>
<td>36.12</td>
</tr>
<tr>
<td>20</td>
<td>.95</td>
<td>.90</td>
<td>.00</td>
<td>.50</td>
<td>.26</td>
<td>20</td>
<td>43.53</td>
<td>35.57</td>
</tr>
<tr>
<td>20</td>
<td>.95</td>
<td>.80</td>
<td>.24</td>
<td>.76</td>
<td>.50</td>
<td>10</td>
<td>32.10</td>
<td>30.06</td>
</tr>
<tr>
<td>20</td>
<td>.95</td>
<td>.80</td>
<td>.24</td>
<td>.76</td>
<td>.50</td>
<td>14</td>
<td>35.67</td>
<td>31.67</td>
</tr>
<tr>
<td>20</td>
<td>.95</td>
<td>.80</td>
<td>.24</td>
<td>.76</td>
<td>.50</td>
<td>18</td>
<td>37.77</td>
<td>33.87</td>
</tr>
<tr>
<td>20</td>
<td>.95</td>
<td>.80</td>
<td>.24</td>
<td>.76</td>
<td>.50</td>
<td>20</td>
<td>38.51</td>
<td>32.75</td>
</tr>
<tr>
<td>40</td>
<td>.95</td>
<td>.95</td>
<td>.24</td>
<td>.76</td>
<td>.50</td>
<td>10</td>
<td>32.79</td>
<td>29.50</td>
</tr>
<tr>
<td>40</td>
<td>.95</td>
<td>.95</td>
<td>.24</td>
<td>.76</td>
<td>.50</td>
<td>14</td>
<td>40.07</td>
<td>34.06</td>
</tr>
<tr>
<td>40</td>
<td>.95</td>
<td>.95</td>
<td>.24</td>
<td>.76</td>
<td>.50</td>
<td>18</td>
<td>45.01</td>
<td>37.60</td>
</tr>
<tr>
<td>40</td>
<td>.95</td>
<td>.95</td>
<td>.24</td>
<td>.76</td>
<td>.50</td>
<td>20</td>
<td>46.86</td>
<td>39.59</td>
</tr>
<tr>
<td>40</td>
<td>.95</td>
<td>.80</td>
<td>.24</td>
<td>.76</td>
<td>.50</td>
<td>10</td>
<td>19.45</td>
<td>18.72</td>
</tr>
<tr>
<td>40</td>
<td>.95</td>
<td>.80</td>
<td>.24</td>
<td>.76</td>
<td>.50</td>
<td>14</td>
<td>21.23</td>
<td>20.90</td>
</tr>
<tr>
<td>40</td>
<td>.95</td>
<td>.80</td>
<td>.24</td>
<td>.76</td>
<td>.50</td>
<td>18</td>
<td>22.38</td>
<td>21.83</td>
</tr>
<tr>
<td>40</td>
<td>.95</td>
<td>.80</td>
<td>.24</td>
<td>.76</td>
<td>.50</td>
<td>20</td>
<td>22.88</td>
<td>22.59</td>
</tr>
<tr>
<td>40</td>
<td>.95</td>
<td>.60</td>
<td>.24</td>
<td>.76</td>
<td>.50</td>
<td>10</td>
<td>14.47</td>
<td>14.42</td>
</tr>
<tr>
<td>40</td>
<td>.95</td>
<td>.60</td>
<td>.24</td>
<td>.76</td>
<td>.50</td>
<td>14</td>
<td>16.00</td>
<td>15.84</td>
</tr>
<tr>
<td>40</td>
<td>.95</td>
<td>.60</td>
<td>.24</td>
<td>.76</td>
<td>.50</td>
<td>18</td>
<td>17.14</td>
<td>16.68</td>
</tr>
<tr>
<td>40</td>
<td>.95</td>
<td>.60</td>
<td>.24</td>
<td>.76</td>
<td>.50</td>
<td>20</td>
<td>17.61</td>
<td>17.34</td>
</tr>
<tr>
<td>20</td>
<td>.95</td>
<td>.60</td>
<td>.24</td>
<td>.76</td>
<td>.50</td>
<td>10</td>
<td>29.53</td>
<td>28.42</td>
</tr>
<tr>
<td>20</td>
<td>.95</td>
<td>.60</td>
<td>.24</td>
<td>.76</td>
<td>.50</td>
<td>14</td>
<td>31.96</td>
<td>29.54</td>
</tr>
<tr>
<td>20</td>
<td>.95</td>
<td>.60</td>
<td>.24</td>
<td>.76</td>
<td>.50</td>
<td>18</td>
<td>33.38</td>
<td>31.38</td>
</tr>
<tr>
<td>20</td>
<td>.95</td>
<td>.60</td>
<td>.24</td>
<td>.76</td>
<td>.50</td>
<td>20</td>
<td>33.92</td>
<td>31.32</td>
</tr>
</tbody>
</table>

REFERENCES