Adaptive Beamforming and Detection Algorithms Based on the Cholesky Decomposition of the Inverse Covariance Matrix

역 공분산 행렬의 Cholesky 분할에 근거한 적응 빔 형성 및 검출 알고리듬

Young C. Park*, Il W. Cha*, Dae H. Youn*
박 영 철*, 차 일 환*, 윤 대 희*

ABSTRACT

The sample matrix inversion(SMI) procedure may suffer from severe computational complexity as well as numerical instability. In this paper, a method which is a numerically more efficient alternative to the SMI procedure is proposed based on the Cholesky decomposition of the inverse covariance matrix. According to this method, adaptive beamforming and detection algorithms are combined in a unified configuration, where the Cholesky factor of the inverse sample covariance matrix is estimated using a Gram Schmidt(GS) processor which directly operates on the secondary inputs. The main feature of this configuration is that neither the covariance matrix nor the Cholesky factor needs to be estimated explicitly, in addition, the arithmetic efficiency of the proposed configuration results from using systolic processing architectures that take advantage of the GS processor. Computer simulations are conducted to show that the proposed configuration shows comparable performance to the theoretical SMI results. Another configuration which does not involve the secondary inputs is also developed to overcome nonhomogeneous environments, and a computationally attractive lattice-GS structure is considered to reduce the computational complexity of the GS structure.

요 약

SMI 방법은 수치적인 불안정성과 아울러 많은 계산량을 갖는다. 본 논문에서는 역 공분산 행렬의 Cholesky 분할을 이용하여 SMI 방법보다 효율적인 방법을 제안한다. 제안한 방법에서는 적응 빔 형상과 검출이 하나의 구조로 실현되며 이에 필요한 역 공분산 행렬의 Cholesky factor는 secondary 입력으로부터 GS 프로세서를 이용하여 추정한다. 제안한 구조의 중요한 특징은 공분산 행렬과 Cholesky factor를 직접 구할 필요가 없다는 점이며, 또한 GS 프로세서의 장점을 이용한 systolic 구조를 사용함으로써 효율적인 계산을 수행할 수 있다. 모의 실험을 통하여 제안한 방법의 성능과 SMI 방법의 성능을 서로 비교하였다. 또한 nonhomogeneous 환경에서 동작하기 위한 방법이 제시되었으며, 아울러 계산량이 많은 GS 구조의 단점을 극복하기 위해 lattice-GS 구조를 이용하는 방법을 제안하였다.

I. Introduction

The objective of an adaptive detector is to detect a target of known form in the presence of noise(interference) which is assumed to be Gaussian, but whose covariance matrix is totally unknown. Brennan and Reed[1] derived a beamformer equation to maximize the probability of detection. However, the exact application of an optimum detection scheme requires a priori knowledge of the noise covariance matrix. Usually in pr-

A.S.S.P. Lab, Dept. Electronic Eng., Yonsei Univ.
접수일자 : 1993. 4. 17

actice, this covariance matrix is not available and must be estimated using an adaptive technique.

In [6], Reed et al. evaluated the performance of an adaptive beamformer, where the sample matrix inversion(SMI) procedure was developed. The SMI procedure uses two sets of data : the primary data, where the known signal may exist and the secondary inputs, which are assumed to contain only noise. An estimate of the noise covariance is directly formed by using independent secondary inputs, which is then used in place of the true covariance in the weight vector equation. To generate the beamformer output, the estimated weight vector is applied to the primary vector. Finally, the output of the beamformer is compared with a threshold for signal detection. The performance of the above procedure was well analyzed in [6]. However, the test statistics in [6] does not provide constant false alarm rate(CFAR) behavior since the tere is no rule to determine the threshold to achieve a given probability of false alarm(Pfa). A generalization of the CFAR detectors was made by Kelly[2]. Kelly derived the generalized likelihood ratio test(GLRT) which is independent of both the level and the structure of the true covariance. Robey et al.[3] also proposed a CFAR detector called the adaptive matched filter(AMF).

Both the GLRT as well as the AMF test necessarily use the SMI procedure to estimate and invert a sample covariance matrix. As noted in [6], the SMI technique converges with the smallest possible number of samples, but requires heavy computational load and suffers from numerical instability. To overcome such problems, Cholesky decomposition technique can be applied[7]. When the dimension of the covariance matrix is large, the Cholesky decomposition technique about half the operations required by direct matrix inversion. Nevertheless, the main difficulty of the Cholesky decomposition approach is that it requires the explicit formation of the sample covariance matrix as in the SMI procedure[7].

One way to further increase the computational and numerical efficiency of the Cholesky decomposition method is to employ Gram-Schmidt (GS) orthogonalization[8]. As a result of the GS orthogonalization of the input data, the inverse covariance matrix can be decomposed into the product of a lower triangular matrix and its complex transpose. This transformation is called the Cholesky decomposition of the inverse covariance matrix[8].

It this paper, an efficient method of implementing the adaptive algorithms which concerns the SMI procedure is developed based on the Cholesky decomposition of the inverse covariance matrix. This is an alternative method of the SMI procedure, but provides numerically more efficient performances than the SMI since it avoids both the estimation as well as the inversion of the sample covariance matrix. According to the developed method, adaptive beamformer and detectors are combined in a unified structure. To estimate the lower triangular matrix(called the Cholesky factor), a GS processor which performs a series of orthogonal projections using the secondary inputs in a systematic manner is comprised in the developed configuration. The main feature of the configuration is that neither the covariance matrix nor the Cholesky factor needs to be estimated explicitly. Instea, parameters estimated in the process of GS orthogonalization are directly applied to two nonadaptive GS processors which operate on the primary data and the steering vector, respectively. In addition, the arithmetic efficiency of the proposed configuration results from using systolic processing architectures that take advantage of the GS processor.

Another configuration implementing adaptive beamforming and detection algorithms based on the primary data is developed to overcome the problem of the limited secondary inputs, which may occur in severely nonhomogeneous environments. Also, a computationally efficient lattice GS structure is considered to reduce the computational complexity of the GS structure.

This paper is organized as follows. In Section

Ⅱ, we briefly review the adaptive array processing algorithms. A configuration unifying the adaptive realization of the beamformer and the detector is described in Section Ⅲ. A solution to the problem of the limited secondary inputs is presented in Section Ⅳ and a computationally attractive lattice-GS structure is applied to the proposed configuration in Section Ⅴ. In Section Ⅵ, we present simulation results. Finally, Section Ⅶ summarizes our results.

Ⅱ. Adaptive Array Processing Algorithms

In an array processor, a data vector is composed of samples from several sensors. Fig.1 illustrates an array processor with N array elements and M taps. Let $\underline{r}^T(k)=[x_1(k)\ x_2(k)\cdots x_N(k)]$ denote a sample vector at time k. The $L(=N\times M)$-dimensional primary data vector which may contain a signal or a target return consists of M sample vectors :

$$\underline{x}^T(k)=[\underline{r}^T(k)\,\underline{r}^T(k-1)\cdots\underline{r}^T(k-M+1)]. \qquad (2.1)$$

The input data is commonly assumed to be a complex Gaussian random vector with mean 0 under hypothesis H0, mean $\alpha\underline{s}$ under hypothesis H1, and covariance matrix $\underline{M}x(k)$. Here, $\underline{s}$ and α denote a known steering vector in the look direction and an unknown complex scalar, respectively.

Now, the detection problem can be written as

$$\underline{x}(k)=\alpha\underline{s}+\underline{n}(k) \quad :H1 \qquad (2.2a)$$

$$\underline{x}(k)=\underline{n}(k) \quad :H0 \qquad (2.2b)$$

where $\underline{n}(k)$ denotes an interference(noise+clutter) vector. For Gaussian distributed interference with known covariance matrix $\underline{M}_x(k)$, Brennan and Reed[1] derived an optimal beamformer maximizing the probability of detection whose weight vector is given by

$$\hat{\underline{w}}=c\,\underline{M}_x^{-1}(k)\underline{s}^*. \qquad (2.3)$$

Maximizing the log likelihood-ratio with respect to the unknown complex amplitude α yields the maximum-likelihood(ML) estimate of α[4,5] :

$$\hat{\alpha}=\frac{\underline{x}^T(k)\,\underline{M}_x^{-1}(k)\underline{s}^*}{\underline{s}^T\underline{M}_x^{-1}(k)\underline{s}^*}, \qquad (2.4)$$

and the generalized likelihood-ratio test(GLRT) is[4,5] :

$$\frac{|\underline{x}^T(k)\,\underline{M}_x^{-1}(k)\underline{s}^*|^2}{\underline{s}^M\underline{X}_x^{-1}(k)\underline{s}^*}\ \underset{H0}{\overset{H1}{\gtrless}}\ \gamma. \qquad (2.5)$$

The main difference between two problems of beamforming and detection is that we are concerned with a decision rule for target presence,

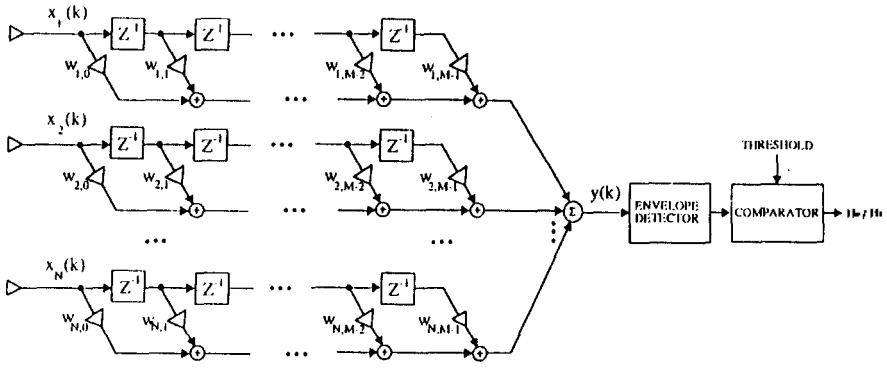


Fig 1. Schematic diagram of an array doppler processor with N elements and M taps.

rather than a filter which will reduce the interference before the data are passed on for further processing and eventually detection[2]. It can be seen from the GLRT equation that the beamforming process accompanied by the nulling of interference seems to be a preliminary step in the detection process. From this point of view, the GLRT can be divided into two processing steps: firstly, a beamformer performing a preliminary processing step produces its output as

$$y(k) = x^T(k)\, w, \tag{2.6}$$

where the weight vector for conventional weighting of the sample vector $\underline{x}(k)$ is given by (2.3). From (2.4) and (2.6), it can be found that $y(k)$ becomes the ML estimate of α when $c = 1/\{\underline{s}^T \underline{M}_x^{-1}(k)\underline{s}^*\}$. Next the beamformer output is passed on to the detector to be tested for target detection. When $c = 1$, the GLRT can be rewritten as

$$|y(k)|^2 \underset{H_0}{\overset{H_1}{\gtrless}} \gamma\delta \tag{2.7}$$

where the threshold γ is multiplied by a scalar constant $\delta = s^T M_x^{-1}(k)s^*$ which is a normalization factor to provide the desired CFAR behavior. If the interference covariance matrix were known, then we would use the optimum detector for target detection. In practice, however, the covariance matrix is unknown and must be estimated by using an adaptive technique.

For practical consideration, it is commonly assumed that there exist a set of secondary data vectors $\underline{x}^1(k), x^2(k), \ldots \underline{x}^K(k)$, which are K independent and identically distributed L-variate complex vectors. The sample covariance matrix based on these secondary inputs is[4]

$$\hat{\underline{M}}_x(k) = \frac{1}{K} \sum_{i=1}^{K} x^{i*}(k) x^{iT}(k). \tag{2.8}$$

Reed et al.[6] have characterized the performance of an adaptive beamformer in which $\hat{M}_x(k)$ is substituted in place of $M_x(k)$ in the beamformer equation.

Kelly[2] derived the GLRT for the problem in which the primary and secondary data vectors were given and both α and $\underline{M}_x(k)$ were unknown parameters. The resulting test is

$$\frac{|x^T(k)\, \hat{\underline{M}}_x^{-1}(k) s^*|^2}{s^T \hat{\underline{M}}_x^{-1}(k) s^* \{1 + (1/K)\, x^T(k) \hat{\underline{M}}_x^{-1}(k) x^*(k)\}} \underset{H_0}{\overset{H_1}{\gtrless}} K\eta. \tag{2.9}$$

Robey et al.[3] also proposed a simplified test statistic that was a limiting case of the GLRT, called the adaptive matched filter(AMF) test:

$$\frac{|\underline{x}^T(k)\, \hat{\underline{M}}_x^{-1}(k) s^*|^2}{s^T \hat{\underline{M}}_x^{-1}(k) s} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma. \tag{2.10}$$

In addition to the constant false alarm rate (CFAR) behavior, both tests in (2.9) and (2.10) have comparable performance to signals aligned with the assumed direction of arrival[3]. As previously stated, the ML estimate of the interference covariance must be formed and then inverted to implement both the GLRT as well as the AMF test.

In the next section, we present a method of implementing the adaptive beamforming and detection algorithms based on the Cholesky decomposition of the inverse covariance matrix.

Ⅲ. A Configuration Unifying the Implementation of Adaptive Beamforming and Detection Algorithms

Under the hypothesis H_0, sample data vector consists of interferences. We can apply the GS orthogonalization to the returned echoes to obtain statistically independent random variables. Consider a GS-transformed vector $\underline{d}^T(k) = [d_0(k)\, d_1(k) \ldots d_{L-1}(k)]$ given by

$$\underline{d}(k) = L_d \underline{x}(k), \tag{3.1}$$

where the GS transformation matrix L_d is a lower triangular matrix whose diagonal elements are all

one[8]. Now $\underline{d}(k)$ is zero-mean Gaussian random variable, and its covariance matrix is a diagonal matrix whose diagonal elements are given by output sample powers :

$$\underline{R}_d(k) = \underline{L}_d^* \underline{M}_x(k) \underline{L}_d^T = \mathrm{diag}\{P_{d,0}(k), P_{d,1}(k), \ldots, P_{d,L-1}(k)\} \quad (3.2)$$

where $P_i(k) = E\{|d_i(k)|^2\}$. Using (3.2), the inverse of $\underline{M}_x(k)$ is decomposed into the product of a lower triangular matrix and its complex transpose, called the Cholesky decomposition[7] :

$$\underline{M}_x^{-1}(k) = \underline{L}_d^T \underline{R}_d^{-1}(k) \underline{L}_d^* = \underline{A}^T(k) \underline{A}^*(k) \quad (3.3)$$

where the Cholesky factor $\underline{A}(k) = \underline{R}_d^{-1/2}(k) \underline{L}_d$, which is also a lower triangular matrix whose diagonal elements are $P_i^{-1/2}(k)$, $0 \le i \le L-1$.

Defining a whitened vector $\underline{x}_a(k) = \underline{A}(k) \underline{x}(k)$, $\underline{x}_a(k)$ is zero-mean Gaussian random variable with covariance matrix equal to $\underline{I}_L$, $(L \times L)$ identity matrix. By inverting the Cholesky factor, we can evaluate the beamformer output :

$$y(k) = \underline{x}^T(k) \underline{M}_x^{-1}(k) \underline{s}^* = (\underline{A}^{-1}(k) \underline{x}_a(k))^T \underline{M}_x^{-1}(k) \underline{s}^* = \underline{x}_a^T(k) \underline{s}_a^* \quad (3.4)$$

where s_a denotes a transformed steering vector : $\underline{s}_a = \underline{A}(k) \underline{s}$. When $c = 1$, we can see that the weight vector in (2.3) can be rewritten as

$$\underline{\hat{w}} = \underline{A}^T(k) \underline{s}_a^*. \quad (3.5)$$

Also, we can evaluate the normalization factor :

$$\delta = \underline{s}^T \underline{M}_x^{-1}(k) \underline{s}^* = \underline{s}_a^T (\underline{A}^T(k))^{-1} \underline{M}_x^{-1}(k) (\underline{A}^*(k))^{-1} \underline{s}_a^* = \underline{s}_a^T \underline{s}_a^* \quad (3.6)$$

Now we can obtain a simple GLRT by substituting the inner products in (2.7). It should be noted that the actual covariance does not appear in (3.4)~(3.6), and thus the test statistics avoids both the formation as well as the inversion of the covariance matrix.

As previously noted, we can reconfigure the beamforming and detection algorithms in simple forms using the Cholesky decomposition technique. The Cholesky factor of the inverse covariance matrix, however, requires the GS tranformation matrix $\underline{L}_d$. Morover, an adaptive technique is required to estimate $\underline{A}(k)$ according to the statistical variation of the secondary inputs(i.e. array environment).

From the assumption that there exist K independent secondary vectors, we can define the $(L \times K)$ secondary data matrix $\underline{X}_L(k)$ as

$$X_L^H(k) = \left[x^1(k)\ \underline{x}^2(k) \ldots \underline{x}^K(k) \right]. \quad (3.7)$$

The ML estimate of interference covariance matrix can be rewritten as

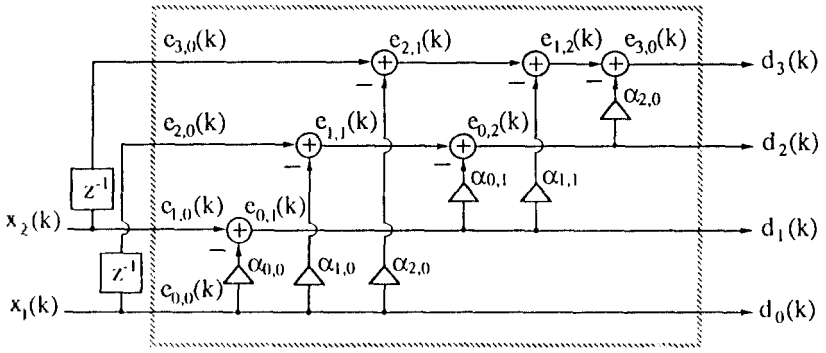
$$\hat{\underline{M}}_x(k) = \frac{1}{K} \underline{X}_L^T(k) \underline{X}_L^*(k). \quad (3.8)$$

After partitioning $\underline{X}_L^H(k)$ as

$$\underline{X}_L^H(k) = \left[\frac{\underline{X}_{L-1}^H(k)}{\underline{y}_{L-1}^T(k)} \right], \quad (3.9)$$

a vector orthogonal to $\underline{X}_{L-1}^H(k)$ is generated using the $(L-1)$th order linear predictor estimating $\underline{y}_{L-1}^H(k)$ using the rest $\underline{X}_{L-1}^H(k)$. We may solve the system of normal equations to estimate a tap weight vector of the $(L-1)$th order predictor. In the process of estimating the tap weight vector the SMI is a compulsory procedure[8].

We can also use the GS structure to generate a vector which is orthogonal to $X_{L-1}^H(k)$. The GS processor produces different sets of mutually uncorrelated random variables in a sequential fashion, and the variable in each set is uncorrelated with those in the preceding set[9]. A multi-channel GS structure is illustrated in Fig. 2 for $M = N = 2$(i.e. $L = 4$). Operations of the GS predictor are based upon elementary processing units. For K inputs per channel, coefficients of the elementary processing units are determined as[10]



MULTICHANNEL GS TRANSFORMER

Fig 2. Schematic diagram of a multi-channel GS processor for M=N=2.

$$\hat{\alpha}_{i,j}(k)=\frac{\sum_{m=0}^{K-1} e_{i-1,j+1}(k-m)\, e_{i-1,j+1}^{*}(k-m)}{\sum_{m=0}^{K-1} |e_{i-1,j}(k-m)|^2}, \quad 1\le i\le N-1,\ 1\le j\le N-i. \tag{3.10}$$

In [10], equivalence between the SMI and GS algorithms has been proved by showing that $\underline{d}_{L-1}(k)$ is orthogonal to $\underline{X}_L(k)$ as the prediction error vector of the $(L-1)$th order linear predictor does[10]. For convenience, this algorithm will be called the maximum-likelihood GS(MLGS).

Recursive equations for the GS structure are summarized as[9]

$$e_{i,j}(k)=e_{i-1,j+1}(k)-\hat{\alpha}_{i,j}\, e_{i-1,j}(k), \quad 1\le i\le N-1,\ 1\le j\le N-i. \tag{3.11}$$

Final output of the GS processor is given by

$$\hat{d}_{L-1}^{T}(k)=\hat{a}_L^{T}(k)\,\underline{X}_L^{H}(k), \tag{3.12}$$

$\hat{a}_{L-1}(k)$ represents the equivalent tap weight[9]:

$$\hat{a}_L^{T}(k)=\underline{u}_L^{T}\,\hat{L}_d(k), \tag{3.13}$$

where $\underline{u}_L^{T}=[0\,0\ldots0\,1]$ denotes a vector picking out the last row in the transformation matrix.

To complete the GS orthogonalization of input data, equivalent tap weight vectors corresponding to successive orders less than $(L-1)$ must be computed. However, the GS processor produces different sets of mutually uncorrelated random variables in a sequential manner, and the variable in each set is uncorrelated with those in the preceding set. Thus, each stage of the GS processor produces uncorrelated output vectors by performing the GS orthogonalization. Defining an $(L\times K)$ transformed data matrix as

$$\underline{D}^{H}(k)=\left[\underline{d}^{1}(k)\,\underline{d}^{2}(k)\cdots\underline{d}^{K}(k)\right]^{T}=\hat{L}_d(k)\,X_L^{H}(k), \tag{3.14}$$

its covariance matrix can be expressed as

$$\frac{1}{K}\,\hat{R}_d(k)=D^{T}(k)\,\underline{D}^{*}(k)=\hat{L}_d^{*}(k)\,\hat{M}_x(k)\,\hat{L}_d^{T}(k) \tag{3.15}$$

which is a diagonal matrix whose diagoal elements are given by the averaged powers of the transformed samples :

$$\hat{P}_{d,i}(k)=\frac{1}{K}\sum_{k=1}^{K}|d_i^{k}|^2, \quad 0\le i\le L-1. \tag{3.16}$$

From (3.15), we can write the inverse of the ML estimate of the clutter covariance as

$$\hat{M}_x^{-1}(k)=\hat{L}_d^{T}(k)\left[\frac{1}{K}\,\hat{R}_d(k)\right]^{-1}\hat{L}_d^{*}(k)$$

$$= \hat{\underline{A}}^{T}(k)\,\hat{\underline{A}}^{*}(k) \tag{3.17}$$

where $\hat{\underline{A}}(k)$ is the desired Cholesky factor of the matrix $\hat{M}_x^{-1}(k)$:

$$\hat{\underline{A}}(k) = \left[\frac{1}{K}\,\hat{R}_d(k)\right]^{-1/2}\hat{L}_d(k). \tag{3.18}$$

Now consider the transformed vectors : $\hat{\underline{x}}_a(k) = \hat{\underline{A}}(k)\underline{x}(k)$ and $\hat{\underline{s}}_a(k) = \hat{\underline{A}}(k)\underline{s}$. We can evaluate the beamformer output

$$y(k) = \hat{\underline{x}}_a^{T}(k)\,\hat{\underline{s}}_a(k), \tag{3.19}$$

and the scalars

$$\hat{\delta}(k) = \hat{\underline{s}}_a^{T}(k)\,\hat{\underline{s}}_a^{*}(k),\quad \hat{\varepsilon}(k) = \hat{\underline{x}}_a^{T}(k)\,\hat{\underline{x}}_a^{*}(k) \tag{3.20}$$

From (3.19) and (3.20), we can see that the beamformer output depends on the primary and steering vectors only through an inner product. When these inner products are substituted into the GLRT in (2.9), we obtain

$$|\hat{y}(k)|^2 \underset{H_0}{\overset{H_1}{\gtrless}} K\gamma\hat{\delta}(k)\{1+(1/K)\hat{\varepsilon}(k)\} \tag{3.21}$$

Also, the AMF test can be expressed in a simple form :

$$|\hat{y}(k)|^2 \underset{H_0}{\overset{H_1}{\gtrless}} \hat{\delta}(k)\gamma \tag{3.22}$$

According to (3.19), (3.21) and (3.22), we can derive a configuration implementing the adaptive beamforming and detection algorithms in a numerically efficient manner as illustrated in Fig. 3. The proposed scheme comprises an adaptive multichannel GS processor which operates on the secondary data vectors to implement the GS orthogonalization. Based on the estimated parameters of the adaptive processor, the Cholesky factor of $\hat{M}_x^{-1}(k)$ is determined as (3.18). The estimated Cholesky factor is, in turn, applied to the primary data vector $\underline{x}(k)$ and the target vector s, respectively. Finally, the inner products in (3.19) and (3.20) are computed to have the beamformer output

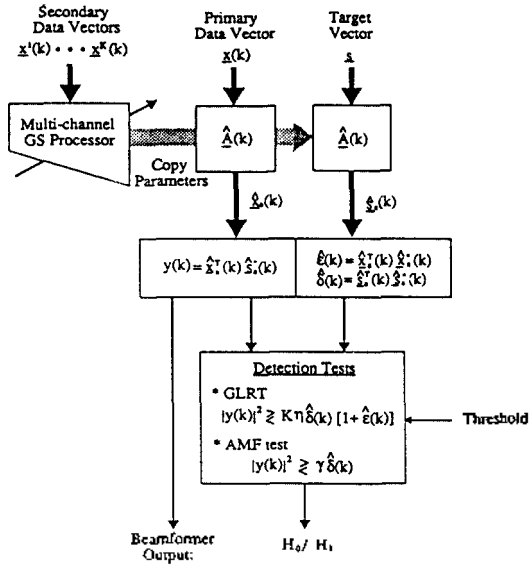


Fig 3. Configuration unifying the implementation of the adaptive beamformer and detectors.

or the test statistics for the detection of a target.

An important feature of the proposed configuration is that neither the sample covariance matrix nor the Cholesky factor needs to be estimated explicitly. Instead, parameters estimated in the process of the GS orthogonalization are directly applied to two nonadaptive GS processors which op-

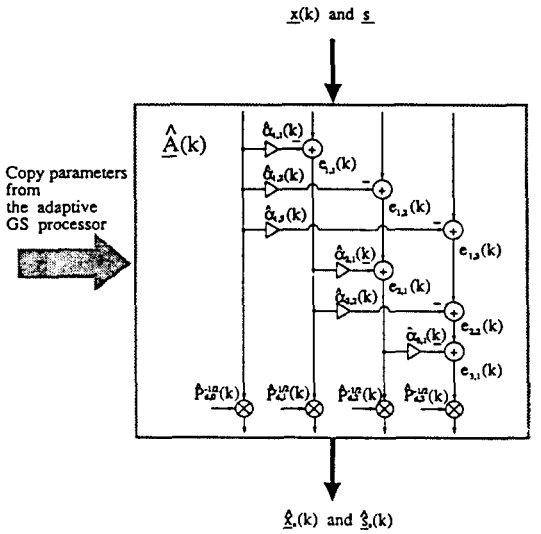


Fig 4. GS processor transforming the input vectors with the estimated Cholesky factor.

erate on the primary data and the steering vector, respectively. Fig. 4 shows the structure of a GS processor transforming input vectors using the estimated Cholesky factor.

Ⅳ. Adaptive Array Processing Based on Primary Data Vectors

A configuration realizing the adaptive beamforming and detection algorithms without the SMI procedure has been derived under the assumption that there exists a set mutually independent secondary vectors. The performance of the SMI technique critically depends on the size of the secondary data set which is very limited in severely nonhomogeneous environments. But if the target signals occur infrequently and do not compete significantly with the incoming noise energy, the interference covariance can be estimated from primary data vectors[6].

Suppose that a separate coherent output is available from each receiving array element. In this case, one approach to the adaptive estimation of the time-varying covariance is to introduce the exponentially time-averaged estimator :

$$\hat{\underline{M}}_x(k) = \lambda \hat{\underline{M}}_x(k-1) + (1-\lambda)\underline{x}^*(k)\underline{x}^T(k) \qquad (4.1)$$

where $0<\lambda<1$ is a forgetting factor.

Applying input sequence to a multichannel GS processor, we get the decorrelated output samples. The best way of estimating the GS transformation matrix from a coherent sequence is the modified GS(MGS) algorithm[11] which performs better than the other orthogonal transformation algorithms such as Householder, and Givens algorithms, as far as numerical accuracy is concerned[7]. Recently, a recursive form of the MGS (RMGS) was proposed[12]. The RMGS performs exact LS estimation and its computational complexity is similar to that of the conventional recursive LS(RLS) algorithm[12].

Using the RMGS algorithm, we can attain the L-dimensional output vector $\underline{q}^T(k) = [q_0(k)\ q_1(k) \cdots q_{L-1}(k)]$, with covariance matrix $\hat{\underline{R}}_q(k)$. $\hat{\underline{R}}_q(k)$ is a diagonal matrix whose diagonal elements are given by exponentially time-averaged estimates :

$$\hat{P}_{q,i}(k) = \lambda \hat{P}_{q,i}(k-1) + (1-\lambda)|q_i(k)|^2, \quad 0 \le i \le L-1. \qquad (4.2)$$

Consequently, the desired Cholesky factor is

$$\hat{\underline{A}}_q(k) = [\hat{\underline{R}}_q(k)]^{-1/2}\,\hat{\underline{L}}_q(k) \qquad (4.3)$$

Now we can obtain the modified beamformer equation and detection tests by replacing $\hat{\underline{A}}_q^T(k)$ $\hat{\underline{A}}_q^*(k)$ in place of $\hat{\underline{M}}_x^{-1}(k)$. From (4.3), we can derive another configuration realizing the adaptive beamforming and detection algorithms in a numerically efficient manner. This configuration does not include the secondary vectors as illustrated in

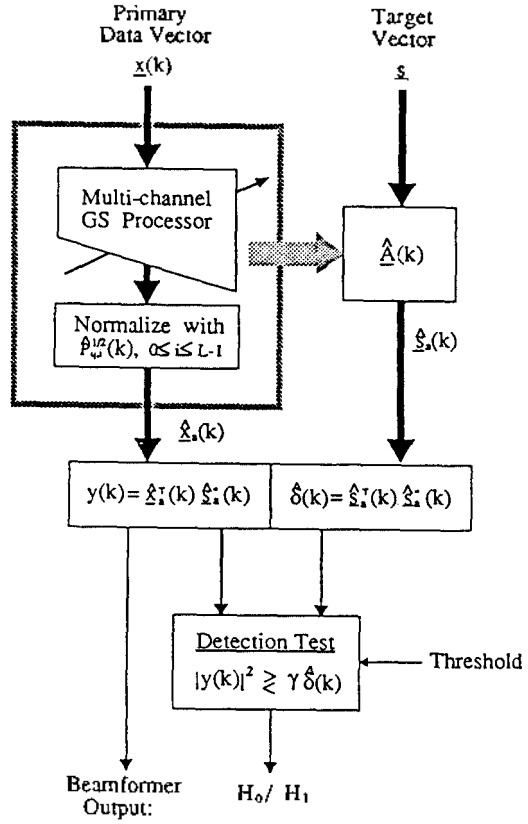


Fig 5. Configuration not concerning the secondary inputs.

Fig. 5. In Fig. 5, an adaptive multi-channel GS processor which operates directly on the primary data vector is included, and the estimated Cholesky factor is then applied to the target vector in the same manner as the configuration in Fig. 3. As previously noted, the transformation of the target vector is performed by using the GS processor instead of estimating the Cholesky factor explicitly.

The RMGS performs the exact LS estimation, however, the computational complexity is similar to that of the RLS algorithm[12]. To remedy this problem, nonexact LS techniques can be used. An example of the nonexact LS method is the least-mean-square(LMS) algorithm. The LMS algorithm[13] is widely used due to its effectiveness as well as its simplicity. The adaptive GS filter whose coefficients are updated using the LMS algorithm has been extensively studied for various applications[8,9]. To improve the slow convergence of the LMS algorithm, the normalized LMS algorithm can be used[8]. The standard gradient(SG) algorithm[13] is a normalize-type LMS algorithm, which was derived to implement the Burg's harmonic-mean algorithm in a recursive way for the LS lattice algorithm. The SG algorithm can be easily applied for the GS processor.

When the LMS algorithm is used, we have an estimate of the inverse covariance matrix :

$$\underline{\hat{S}}(k) = \underline{\hat{L}}_q^T(k)\,\underline{\hat{D}}^{-1}(k)\,\underline{\hat{L}}_q^*(k) = \underline{\hat{A}}_q^T(k)\,\underline{\hat{A}}_q^*(k) \tag{4.4}$$

where

$$\underline{\hat{D}}(k) = \mathrm{diag}\{\hat{P}_{q,0}(k), \cdots, \hat{P}_{q,L-1}(k)\} \tag{4.5}$$

and $\hat{P}_{q,i}(k)$, $0 \le i \le L-1$, are exponentially time-averaged power estimates of $q_i(k)$. From (4.4), we get an estimate of the Cholesky factor :

$$\underline{\hat{A}}_q^T(k) = \underline{\hat{D}}^{-1/2}(k)\,\underline{\hat{L}}_q(k). \tag{4.6}$$

We may use the estimate $\underline{\hat{S}}(k)$ in place of $\underline{\hat{M}}_x^{-1}(k)$ in both the beamformer equation and the detection tests. Regardless of the simplicity of the LMS algorithm, the main drawback is that its convergence speed depends on the eigenvalue spread ratio of the reference input covariance [13]. Thus, it does not guarantee convergence with a small number of data samples, and the orthogonality between the output error and the reference input is satisfied only after the weights converge to their optimal values in contrast to the RMGS algorithm.

V. Consideration of the Lattice-GS Structure

The GS orthogonalization procedure can be implemented using a lattice or a GS structure. The lattice structure has been widely used for various applications[8]. In a multi-channel lattice structure, it is known[8] that the backward prediction error vectors produced at the different stages are mutually orthogonal. However, the elements of an error vector are still correlated. Thus, the covariance matrix of the prediction error vector is a block diagonal matrix. It has been reported[9] that the predictor having the GS structure outperforms the lattice counterpart, since the GS predictor completely orthogonaizes the reference signals. Although complete orthogonalization can be achieved with the GS structure, a maior shortcoming of applying the GS structure to the proposed configuration is the heavy computational load which increases at the rate of $(MN)^2$.

As an effort of overcoming this problem, the multi-channel lattice-GS hybrid structure[15,16] was proposed. The multi-channel lattice-GS structure was attained by using the GS linear combiner in place of the linear combiner summing the weighted elements of the error vectors of the multi-channel lattice structure[15,16]. In the lattice-GS structure, the prediction error vectors are generated by using the lattice structure and the elements of the error vectors are decorrelated by using the GS structure. Fig. 6 shows the i-th stage of the lattice-GS structure.

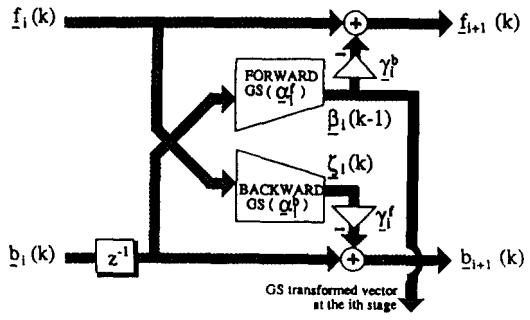


Fig 6. Block diagram of the i-th stage of the lattice-GS structure.

Order updates for the lattice-GS structure are summarized as

$$\underline{f}_{i+1}(k) = \underline{f}_i(k) - \gamma_{i+1}^b \underline{\beta}_i(k-1), \quad (5.1a)$$

$$\underline{b}_{i+1}(k) = \underline{b}_i(k-1) - \gamma_{i+1}^{f\,H} \underline{\zeta}_i(k), \quad 1 \le i \le M, \quad (5.1b)$$

where $\underline{f}_i(k)$ and $\underline{b}_i(k)$ are the $(N \times 1)$ forward and backward prediction error vectors at the i th stage, respectively, $\underline{\zeta}_i(k)$ and $\underline{\beta}_i(k)$ are the GS-transformed versions of $\underline{f}_i(k)$ and $\underline{b}_i(k)$. In this case, partial correlations are defined as

$$\gamma_{i+1}^b(k) = E[\underline{f}_i^*(k)\, \underline{\beta}_i^T(k)], \quad (5.2a)$$

$$\gamma_{i+1}^f(k) = E[\underline{\zeta}_i^*(k)\, \underline{b}_i^T(k)]. \quad (5.2b)$$

Now the covariance matrix of $\underline{\beta}_i(k)$ is given by

$$\underline{R}_\beta(k) = \underline{L}_\beta^*(k)\, \underline{M}_x(k)\, \underline{L}_\beta^T(k). \quad (5.3)$$

which is naturally a diagonal matrix. We can use (5.3) to obtain the Cholesky factor of the inverse of $\underline{M}_x(k)$ as following :

$$\underline{A}(k) = \underline{R}_\beta^{-1/2}(k)\, \underline{L}_\beta(k) \quad (5.4)$$

To implement the i-th stage of the multi-channel lattice-GS structure, $3N^2$ coefficients are required ; this is a big save compared with what we have to pay to completely decorrelate the reference signals using only the GS structure[16].

Total number of coefficients required to implement the multi-channel predictor using the GS and the lattice-GS structures can be summarized as

GS : $MN(MN-1)/2$
lattice−GS : $3N(N-1)(M-1)+N(N-1)/2$.

Fig. 7 shows the total number of coefficients of a GS and a lattice GS structures for $N=6$. We can see that the computational effectivess of the lattice GS structure increasee in proportion to the increase of the taps. Thus, computational effectiveness is ensured when the lattice-GS structure is applied to the proposed configurations in Fig. 3 and Fig. 5 in place of the GS structure. This effectiveness is emphasized by the increase of the array dimension.

When the coefficients of the lattice GS predictor are determined in the exact LS sense based on the K independent secondary data vectors, the inverse of the sample covariance matrix can be partitioned into the product of the Cholesky factor and its complex transpose. To estimate the Cholesky factor using the lattice-GS structure, it is required to combine the multi-channel LS lattice and RMGS algorithms[15]. But gradient-type algorithms such as the LMS and the SG can be used to have a more computationally attractive configuration.

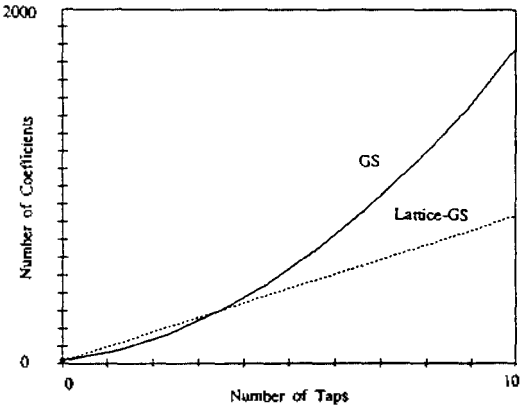


Fig 7. Total number of coefficients of the multi-channel GS and lattice-GS structure for N=6.

VI. Simulation Results and Discussion

In this section, performances of the proposed approach are evaluated and compared with the theoretical SMI results via computer simulations.

For experimental purposes, it was assumed that the array had six sensor elements on a line spaced with the distance of half-wavelength and each element had three taps (MN=18). The environment had 12 independent interference sources which were restricted to two angular regions, $-25^\circ<\theta<-40^\circ$ and $40^\circ<\theta<50^\circ$. In the look direction two independent interferers were included. Interference signals were presumed to have the Gaussian-shaped frequency spectrum. Gaussian interferences for computer simulations were generated using the discrete Fourier transform(DFT) method[17]. Parameters assumed in the simulations are listed in Table 1.

Table 1. Interference parameters for computer simulations.

N=5, M=3
$d/\lambda=0.5$ where d=distance between the adjacentsensors.
Omnidirectional sensor patterns.
2 interference sources in look direction :
$f_{o1}=0$ $f_{o2}=0.1fs$
$\rho_1=0.99$ $\rho_2=0.98$
$CNR_1=40dB$ $CNR_2=40dB$
4 interference sources are equally spaced in the region $40^\circ<\theta<50^\circ$
6 interference sources are equally spaced in the region $-25^\circ<\theta<-40^\circ$
All interference sources in the sidelobe region have the same parameters :
$f_o=0.3fs$
$\rho=0.95$
CNR=40dB

The steady-state covariance matrix $\underline{M}_x$ was first computed from 4,000 independent secondary data vectors. From this, optimum weight vector of the GLRT detector was obtained from (2.3). This weight, in turn, were then used to compute the steady-state signal-to-noise ratio(SNR) to be 22.8 dB. To compare initial convergence behaviors, we computed the instantaneous output SNR values defined as

$$SNR(k)\,dB = 10\log_{10}\frac{|\underline{\hat{s}}^T\underline{w}(k)|^2}{\underline{\hat{w}}^H(k)\,\underline{\hat{M}}_x\,\underline{\hat{w}}(k)}. \qquad (6.1)$$

where the weight vector $\underline{\hat{w}}(k)$ was computed from the estimated Cholesky factor :

$$\underline{\hat{w}}(k)=\hat{A}^T(k)\,\hat{\alpha}^*(k). \qquad (6.2)$$

Initial values of the output SNR of the proposed method and the theoretical SMI curve are illustrated in Fig. 8, where the theoretical SMI curve was obtained from the SMI theory[5]. Corresponding plots of SNR values were obtained by ensemble averaging over 20 independent trials of simulations. For the SMI and RMGS algorithms, number of samples implies the number of secondary vectors with which the covariances are averaged. More secondary vectors imply an accurate estimation of the matrix and reduced SNR loss accompanied by a longer transient time[5]. For the RMGS, SG and the LMS algorithms, the number of samples denotes the number of independent data vectors(i.e. the number of primary vectors).

The exact LS methods(MLGS and RMGS) show a rapid convergence rate to the steady-state SNR value of 22.8 dB and the comparable performance to the theoretical SMI. Also, as expected, the proposed configuration shows the same results as the SMI algorithm when the MLGS algorithm is used. The SG algorithm shows slower convergence speed than the case of the exact LS algorithms. The LMS algorithm shows the worst performance in terms of the convergence speed, however, simplicity and effectiveness are the main advantages of the LMS algorithm. It was observed that the exponential window methods(SG and RMGS) were accompanied by the steady-state variations of SNR values caused by the exponential weighting factor λ which was introduced to ensure that the data in the distant past were forgotton in order to track the changing environment. It is known that the steady-state and

transient performance can be trade off in several ways according to different parameters to be used. In the simulations, 0.002 and 0.98 were used for μ and λ values, respectively. Those values were experimentally determined to have the same steady-state variances.

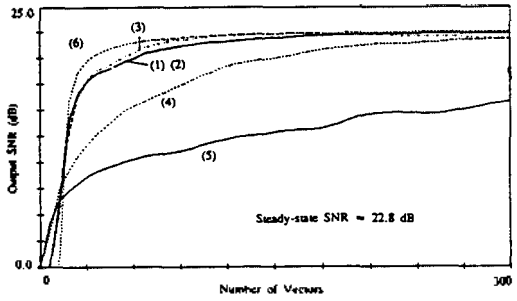


Fig 8. Output SNR values of the proposed configuration for different algorithms : (1)SMI, (2) MLGS, (3)RMGS with λ=0.98, (4)SG with λ =0.98, (5)LMS with μ=0.02 and (6)SMI in theory.

Performances of the proposed configurations were compared for different algorithms in terms of beampattern and frequency response. Fig. 9 and Fig. 10 show beampatterns at the frequency of 900 Hz and frequency responses of the proposed scheme, respectively. Results were evaluated for k=100. From Fig. 9, it can be seen that sidelobes are reduced in angular intervals $-25°<\theta<-40°$ and $40°<\theta<50°$ where scatters are located. We can also see that the exact LS methods perform better than the nonexact methods. Similar results can be observed from Fig. 10 where both clutter sources in look direction are correctly rejected for all the algorithms.

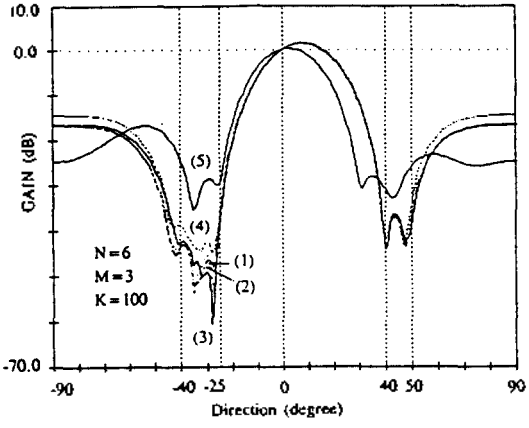


Fig 9. Beampatterns of the proposed configuration for different algorithms : (1)SMI, (2)MLGS, (3) RMGS, (4)SG and (5)LMS.

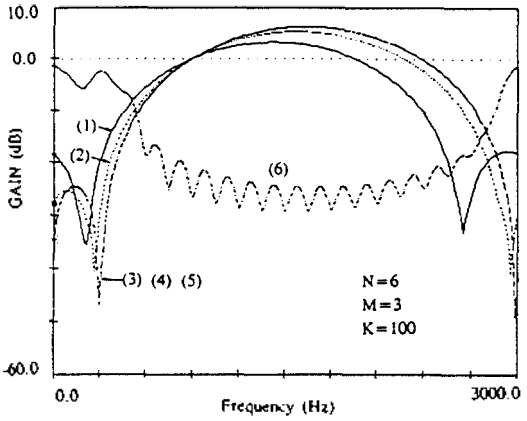


Fig 10. Frequency responses of the proposed configuration in the look direction. (1)SMI, (2)MLGS, (3)RMGS, (4)SG and (5)LMS. (6)shows the spectrum of the background clutter.

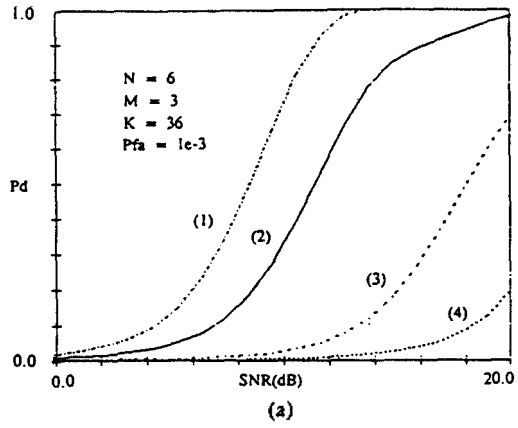
For each adapted weight vector, experimental detection probability(Pd) versus SNR curves are generated by evaluating the Q-function[1] :

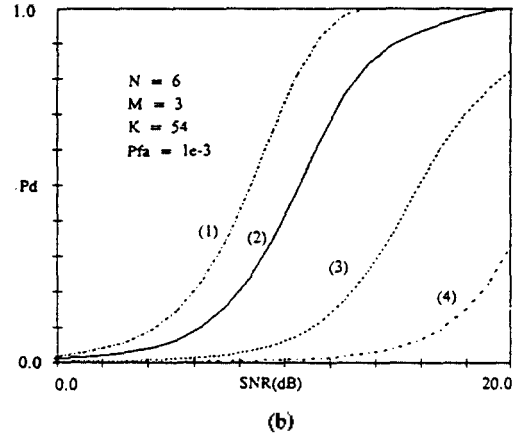
$$P_d = Q\left[\frac{|\alpha|^2 |\hat{\underline{w}}^T(k)\,\underline{s}^*|^2}{\underline{w}(k)\,\hat{\underline{M}}_x(k)\,\hat{\underline{w}}^*(k)},\ 2\log\frac{1}{Pfa}\right] \quad (6.3)$$

where Pfa was set to a constant and α was chosen to give some desired value of SNR according to

$$SNR = |\alpha|^2 \underline{s}^T \hat{\underline{M}}_x^{-1}(k)\,\underline{s}^*. \quad (6.4)$$

In detection tests, we examined the experimental Pd for SNR values in the range of 0 to 20 dB in 1-dB increment. To study the effect of the sample size K on the detector performance, we considered the case K=36, 54, 72. The resulting Pd versus SNR curves are shown in Fig. 11 and Fig.





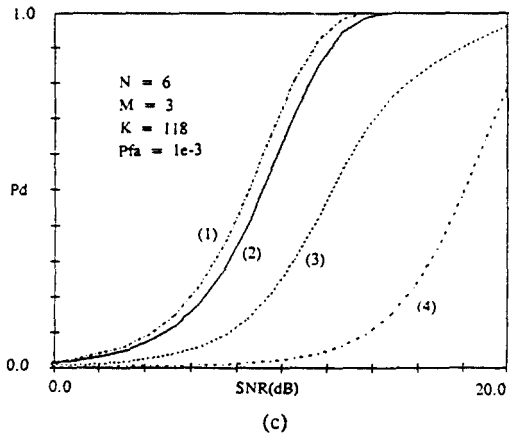
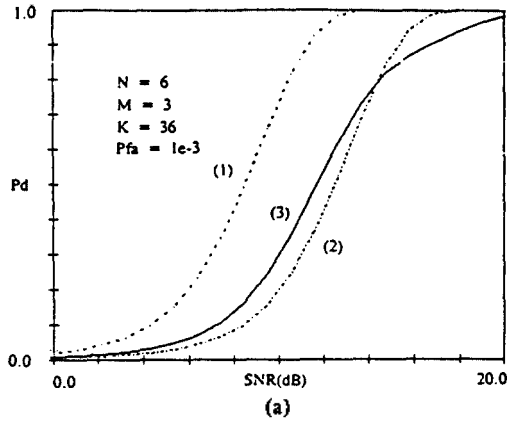
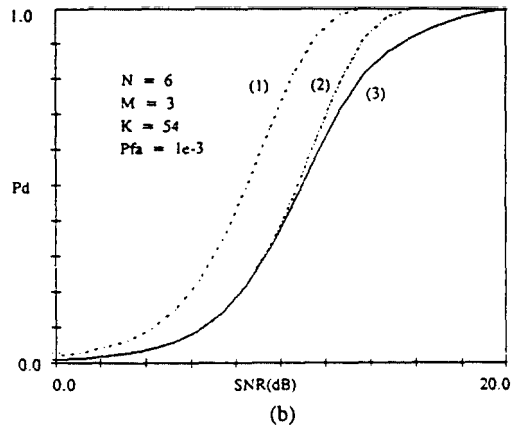


Fig 11. Pd versus SNR for different sample numbers. (1)MF, (2)AMF in theory and (3)proposed configuration.





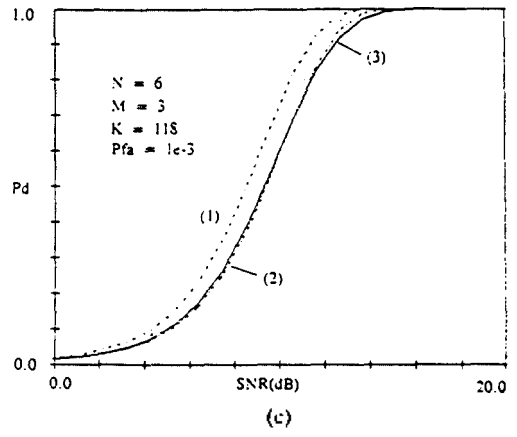


Fig 12. Pd versus SNR of the configuration in Fig. 5 for different algorithms : RMGS(2), SG(3) and LMS(4). Pd values of the MF is shown(1).

12, where Pd of the matched filter(MF) versus the SNR curve evaluated at $Pfa = 1.0 \times 10^{-3}$ and the theoretical Pd of the AMF detector versus SNR curve[3] are given. Considering Figs. 11(a)-(c), the proposed configuration shows a comparable bahavior to the theoretical AMF detector in the detection performance. The SNR loss relative to the MF detector improved in accordance with the increase of K and the detection performance degrades systematically with decreasing K as well. Figs. 12(a)-(c) show the Pd versus SNR curves for the configuration in Fig. 5, where the RMGS, SG and LMS algorithms are considered. The detection performances are all quite similar to the results in Fig. 8. The LMS algorithm shows poor results since the orthogonality between the output error and the reference input is not satisfied until the sample index reaches to corresponding K values in contrast to the RMGS algorithm.

Finally, performances of the lattice-GS structure are demonstrated. Simulations were perfor med with the same parameters listed in Table 1. Initial convergence behaviors of the GS and lattice-GS predictors were compared in terms of instantaneous output SIR values. Fig. 13 shows output SNR curves of two approaches when the SG algorithm is used. Although the nonexact method(SG algorithm) is used, we can see that performances of the lattice-GS predictor is comparable to the results of the GS structure. Similar results can be observed from beampatterns and frequency responses shown in Fig. 14 and Fig. 15, respectively.

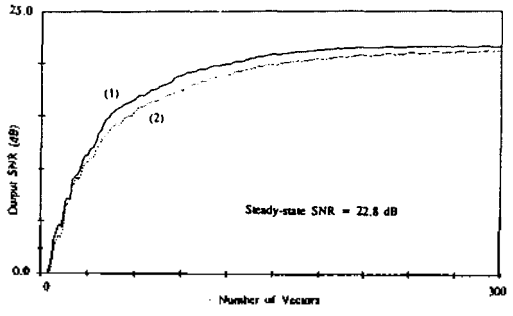


Fig 13. Output SNR values of the proposed configuration for the case when the SG algorithm is used to compute the coefficients of the GS(1) and lattice-GS(2) structures.

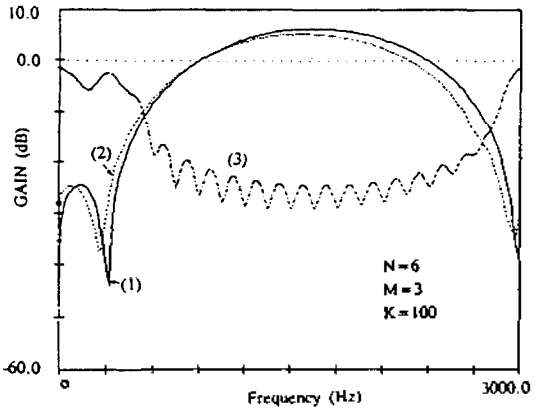


Fig 14. Beampatterns of the proposed configuration for the cases when the GS(1) and lattice-GS(2) structures are considered.

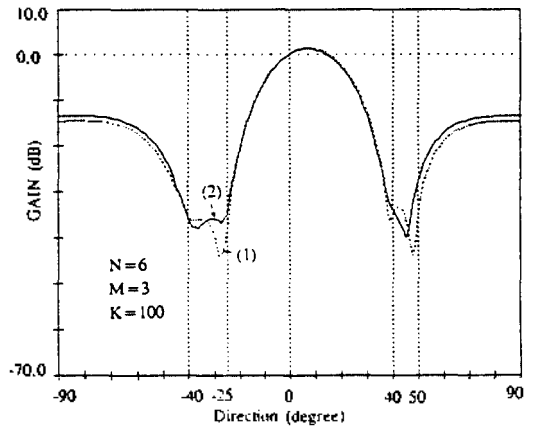


Fig 15. Frequency responses of the proposed configuration for the cases when the GS(1) and lattice-GS(2) structures are considered. (3) shows the spectrum of the background clutter.

Ⅵ. Summary

An efficient method of implementing the SMI procedure has been developed based on the Cholesky decomposition of the inverse covariance matrix. According to this method, we have com-

bined adaptive beamforming and detection algorithms in a unified configuration. The proposed configuration is a numerically more efficient alternative to the SMI algorithm, but does not require the explicit estimation of either the sample covariance matrix or the Cholesky factor.

To obtain the Cholesky factor of the inverse covariance matrix, a multi-channel adaptive GS processor which operates on the secondary inputs is comprised in the configuration. Instead of forming the Choelsky factor explicitly, parameters estimated in the process of the GS orthogonalization are directly applied to two nonadaptive GS processors which operate on the primary data and the steering vector, respectively. In this way, the input and target vectors are transformed by the Cholesky factor. Thus, the proposed configuration has the advantage of systolic processing architectures which ensure the arithmetic efficiency. It has been shown that the proposed method provides a very fast rate of convergence and comparable performance to theoretical SMI results when algorithms computing the exact LS solution were used.

We have developed another configuration, where the secondary inputs have not been included, to overcome nonhomogeneous environments. This is a very useful configuration especially for the case when the target signals are encountered infrequently and do not compete significantly with the incoming energy. Also, we have applied a computationally attractive lattice-GS structure to the proposed configurations. A multichannel lattice-GS structure requires $3N^2$ coefficients to implement its ith stage, which is a big saving compared with what we have to pay to completely decorrelate the reference signals using only the GS structure in which the computational load increases at the rate of $(MN)^2$.

REFERENCES

1. L. E. Brennan, and I. S. Reed, "Theory of Adaptive Radar," IEEE Trans. Aerospace & Electronic System, Vol AES-9, No.2, pp.237-252, March 1973.
2. E. J. Kelly, "An Adaptive Detection Algorithm," IEEE Trans. Aerospace & Electronic System, Vol AES-22, No.2, pp.115-127, March 1986.
3. F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR Adaptive Matched Filter," IEEE Trans. Aerospace & Electronic System, Vol AES 28, No.1, pp.208-216, Jan. 1992.
4. H. L. Van Trees, Detection, Estimation, and Modulation Theory, New York : Wiley, 1986.
5. L. L. Scharf, Statistical Signal Processing, Addison-Wesley Publishing Co., 1991.
6. I. S. Reed, J. D. Mallett, and L. E. Brennan, "Rapid Convergence Rate in Adaptive Arrays," IEEE Trans. Aerospace & Electronic System, Vil AES-10, No.6, pp.853 863, Nov. 1973.
7. S. M. Yuen, "Exact Least Squares Adaptive Beamforming Using an Orthogonalization Network," IEEE Trans. Aerospace & Electronic System, Vol AES-27, No.2, pp.311-330, March 1991.
8. S. Haykin, Adaptive Filter Theory, Englewood Cliffs, NJ, Prentice-Hall Inc., 1991.
9. N. Ahmed and D. H. Youn "On a Realization and Related Algorithm for Adaptive Prediction," IEEE Trans. Acoust., Speech, Signal Proc., Vol.ASSP-28, No.5, pp.493-497, Oct. 1980.
10. K. Gerlach and F. F. Kretschmer Jr., "Convergence Properties of Gram-Schmidt and SMI Adaptive Algorithms," IEEE Trans. Aerospace & Electronic System, Vol AES-26, No.1, pp.44-56, Jan. 1990.
11. G. H. Golub and C. F. Van Loan, Matrix Computations, Johns Hopkins University Press, Baltimore, 1983.
12. F. Ling, D. Manolakis, and J. G. Proakis, "A Recursive Modified Gram-Schmidt Algorithm for Least-Squares Estimation," IEEE Trans. Acoust., Speech, Signal Proc., Vol.ASSP 34, No.4, pp.829-835, Aug. 1986.
13. B. Widrow et al., "Adaptive Noise Cancelling : Principles and Applications," Proc. IEEE, Vol.55, No.12, pp.2143 2159, Dec. 1975.
14. S. Haykin, "Radar Signal Processing," IEEE Acoust., Speech, Signal Proc., Magazine, pp.2 18, April 1985.
15. F. Ling and J. G. Proakis, "A Generalized Multichannel Least Squares Lattice Algorithm Based on Sequential Processing Stages," IEEE Trans. A-

coust., Speech, Signal Proc., Vol. ASSP-32, No.2, pp.381-389, April 1984.
16. K. M. Kim, Y. C. Park, I. W. Cha and D. H. Youn, "Adaptive Multichannel Lattice-Escalator Filter Structure : An Application to Generalized Sidelobe Canceller," IEEE Trans. Signal Proc., Vol.40, No. 7, pp.1817-1819, July 1992.
17. R. L. Mitchell, Radar Signal Simulation, Dedham, Mass., Artech House, Inc., 1976.

This study was supported in part by research grant from the Agency for Defence Development of Korea.
본 연구의 일부는 국방과학연구소의 연구비 지원에 의해 이루어진 것임

▲Young-Cheol Park

Young-Cheol Park was born in Korea on Feburary 28, 1964. He received B.S., M.S and Ph.D degree in electronic engineering from Yonsel university in 1986, 1988 and 1993. He is working in post doctorial course at Pennsylyania state university. His research interests are in array signal processing adaptive signal processing and its applicat ions.

▲Il-Whan Cha

Il Whan Cha for a photograph and biography, see pp. 52 of the December 1990 issue of this journal.

▲Dae-Hee Youn

Dae-Hee Youn sfor a photograph and biography, see pp. 52 of the December 1990 issue of this journal.