Adaptive Multiple-Band CFAR Detection of an Optical Pattern with Unknown Spectral Distribution

IRVING S. REED, FELLOW, IEEE, AND XIAOLI YU

Abstract—The constant false alarm rate (CFAR) detection algorithm considered by Chen and Reed is generalized to a test which is able to detect the presence of a known optical signal pattern which has nonnegligible unknown relative intensities in several signal-plus-noise bands or channels. This new test and its statistics are analytically evaluated and the signal-to-noise ratio (SNR) performance improvement is analyzed. Both theoretical and computer simulation results show that the SNR improvement factor of this new algorithm using multiple band scenes over the single scene of maximum SNR can be substantial.

The SNR gain of this new detection algorithm and the one given by Chen and Reed are compared. It illustrates that the GSNR of the test using the full data array is always greater than that of using a partial data array. The data base used to simulate this new adaptive CFAR test is obtained from actual LANDSAT image scenes. The present results for optical detection are extendable to radar target detection and to other related detection problems.

I. Introduction

In [1] an adaptive constant false alarm rate (CFAR) detection algorithm was developed from the generalized likelihood ratio (GLR) test. This test criterion in turn is based on the suggestion that most optical image clutter can be modeled as a Gaussian random process with possibly a rapidly fluctuating space-varying mean and a more slowly varying covariance [4], [5].

Such a CFAR test is closely related to the test developed by Kelly in [10] for detecting radar targets. The probability of a false alarm (PFA) of the CFAR test in [1] is a function only of N, the number of samples, and K, the number of reference image scenes used. The CFAR detection algorithm in [1] allows one to find a detection threshold that achieves a fixed PFA which is invariant to intensity changes in the noise background.

The CFAR detection algorithm considered in [1] is suitable only for detecting a target pattern in one main image scene and a number of other noise-only reference image scenes that contain negligible signal energy. However, in many applications, one needs to test for the presence of a signal pattern which has nonnegligible unknown relative intensities in several optical bands. As a consequence it is of importance to generalize the previous CFAR detection algorithm [1] to a test which is able to detect the presence of an optical signal pattern with nonzero intensity in several signal-plus-noise bands or channels. An effort was made in [6] to find and compute the statistics of this generalized CFAR test, but the results were incomplete.

In this paper the approach first considered in [6] to find this new CFAR test is improved and solved. In Section II the general hypothesis test is formulated and found in terms of the GLR principle. The result is a CFAR test for a signal with unknown relative intensities in J channels. If $J = 1$, the resulting test reduces to a well-known CFAR test for one receiving channel. The detection statistic found for this new test is similar to the adaptive array test for spread spectrum communications obtained by Brennan and Reed in [7] except that their test was not obtained from a hypothesis test. The test in [7] for automatic synchronization was derived as a least mean-square criterion.

As noted above, the GLR test under the assumption of unknown background clutter statistics has the CFAR property that the probability of a false alarm (PFA) or its equivalent, the probability of signal detection, given the null or clutter-noise only hypothesis H_0, is independent of the actual covariance matrix of the data. In the statistical literature under the null hypothesis, the GLR criterion as derived here is related to the so-called generalized T^2 statistic, e.g., see [8], [9]. Also, much of the statistical machinery, used here for optical CFAR detection on clutter noise, appears under the general heading of multivariate analysis of variance (MANOVA). However, usually these statistical methods are applied only to obtain the probability of the null hypothesis which is equal to $1 - \mathrm{PFA}$ for the present problem. The explicit results, obtained here using a methodology similar to MANOVA, are specific to the assumed optical model and as such are new to optical detection theory.

In order to analyze the performance of this new CFAR test, the probability density function of the GLR test for both hypotheses is analytically evaluated in Section III. These probability densities are used then to calculate the PFA and the PD as a function of the detection threshold in a manner similar to that utilized in [1]. The PFA of the test is computed in a closed formula which is independent of the covariance matrix of the actual residual clutter noise

Manuscript received February 1, 1989; revised September 7, 1989. This work was supported in part by the Naval Ocean Systems Center under Contract N66001-87-D-0136 and by the Rome Air Development Center (RADC) under Contract F19628-86-K-0008.

The authors are with the Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089-0271.

IEEE Log Number 9037817.

encountered. In Section IV the signal-to-noise ratio (SNR) performance improvement is analyzed. Both theoretical and computer simulation results show that the SNR improvement factor of this new algorithm using multiple band scenes over the single scene of maximum SNR can be substantial. Also, the SNR gain of this new detection algorithm and the one given in [1] are compared. The data base used to simulate this new adaptive CFAR test is obtained from actual LANDSAT image scenes. The probability density obtained here for a real optical signal in residual clutter noise is similar to that found in [7] for complex communication channels. However, the method of derivation is different and extendable to other more general detection problems.

II. Formulation of the Problem

Most statistical models for optical clutter images try to assume a Gaussian probability density function even though such behavior rarely occurs. As an illustration, Fig. 2 shows the intensity histogram of a 32×32 subimage taken from the IR image of the Santa Monica mountains which is shown in Fig. 1. Since a Gaussian distribution model is mathematically tractable to multivariate statistical methods, it is desirable to force Gaussianity on an image by assuming the clutter has a nonstationary mean value as suggested more completely in [2] and [4].

In both [1] and [3], an image of 512×512 pixels is first divided into 256 small 32×32 subimages. It is assumed that the mean can change substantially in the subimage but that the covariance and the cross covariances of the subimages of two or more scenes is slowly varying over the image.

A nonstationary local mean of an image X is obtained from

$$\bar{X} = \frac{1}{L^2}[X * W]$$

where W is an $L \times L$ all-ones matrix or window, and the asterisk (*) denotes convolution. In [3] and [4] it is shown that one simple method to find the window size L is to minimize the third moment of $X^0 = X - \bar{X}$. Fig. 3 shows the histogram of such a residual image X^0. Since a Gaussian distribution has a zero third moment, this minimization tends to create a distribution which is as close to Gaussian as possible. Evidently, this residual image is approximately Gaussian.

The present detection problem is formulated in a manner similar to that used in [1]. First let the column vectors

$$x(n) = [x_1(n), x_2(n), \cdots, x_J(n)]^T \tag{1a}$$

for $n = 1, 2, \cdots, N$ represent the J correlated subimage scenes which may contain an optical signal with known shape and unknown position where it is assumed that $N > J$. Next let

$$S = [s(1), s(2), \cdots, s(N)] \tag{1b}$$

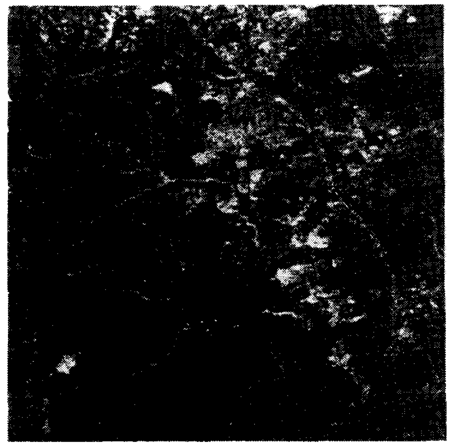


Fig. 1. IR false-color image.

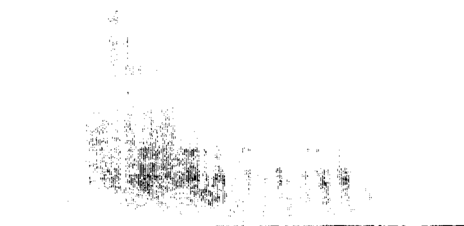


Fig. 2. Histogram of subimage in Fig. 1.

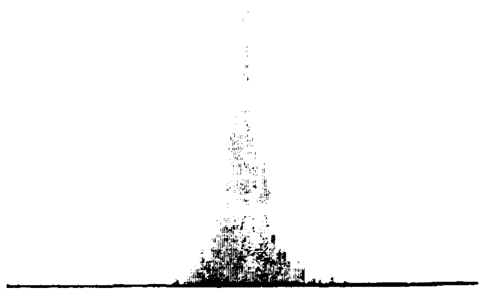


Fig. 3. Histogram of the residual subimage.

be the signal pattern as an N-element row vector, and

$$\boldsymbol{b} = [b_1, b_2, \cdots, b_J]^T \tag{1c}$$

be a J vector of unknown signal intensities corresponding to the J scenes or channels, respectively. The two hypotheses which the adaptive detector must distinguish are given by

$$H_0: \boldsymbol{x}(n) = \boldsymbol{x}^0(n)$$
$$H_1: \boldsymbol{x}(n) = \boldsymbol{x}^0(n) + \boldsymbol{b}s(n) \tag{2}$$

for $n = 1, 2, \cdots, N$ where $\boldsymbol{x}(n)^0$ is the vector of residual clutter noise-only processes.

III. The Generalized Maximum Likelihood Ratio (MLR) Test

In order to distinguish the two hypotheses in (2) the generalized maximum likelihood ratio GLR test, originally due to Neyman and Pearson [16], is used. The GLR principle is best described by a likelihood ratio defined on some sample space X with a parameter set Ω.

It was demonstrated previously [3] (see also [1, appendix A]), that after the appropriate subtraction of a nonstationary mean, the covariance matrix of many optical and IR images can be approximated by diagonal matrices. This indicates that such residual clutter noise is approximately independent and Gaussian from pixel to pixel [1]-[4]. Therefore, it is reasonable to assume that the residual clutter is a zero-mean Gaussian process and independent from spatial sample to sample. For the limitations on this assumption see [1, appendix A] and [4].

Let

$$M \triangleq \boldsymbol{E}\left\{[\boldsymbol{x}(n) - \boldsymbol{E}\boldsymbol{x}(n)][\boldsymbol{x}(n) - \boldsymbol{E}\boldsymbol{x}(n)]^T\right\}$$

be the unknown covariance matrix of random vector $\boldsymbol{x}(n)$ in (1a) for $n = 1, 2, \cdots, N$, then for the current problem

$$\Omega \triangleq \{\theta\} = \left\{[\boldsymbol{b}, M] \mid M > 0\right\}$$

and the likelihood function is

$$\begin{aligned} L(\boldsymbol{b}, M) &= \frac{1}{2\pi^{NJ/2}|M|^{N/2}} \\ &\quad \cdot \exp\left[\frac{-1}{2}\sum_{n=1}^{N}(\boldsymbol{x}(n) - \boldsymbol{E}\boldsymbol{x}(n))^T \right. \\ &\quad \left. \cdot M^{-1}(\boldsymbol{x}(n) - \boldsymbol{E}\boldsymbol{x}(n))\right] \\ &= \frac{1}{2\pi^{NJ/2}|M|^{N/2}} \\ &\quad \cdot \exp\left\{\frac{-1}{2}\operatorname{Tr}\left[M^{-1}(X - \boldsymbol{E}X)(X - \boldsymbol{E}X)^T\right]\right\} \end{aligned} \tag{3}$$

where $|M| \neq 0$ is the determinant of M and

$$X = [\boldsymbol{x}(1), \cdots, \boldsymbol{x}(N)]$$

is a $J \times N$ matrix of vector data $\boldsymbol{x}(n)$ and Tr denotes the matrix trace function.

Next let ω be the region in the parameter space Ω specified by the H_0 hypothesis, then in terms of the complementary subsets ω and $\Omega - \omega$ of Ω define the alternative hypotheses H_0 and H_1 as follows:

$$H_0 \equiv [\boldsymbol{b}, M] \in \omega \quad \text{and} \quad H_1 \equiv [\boldsymbol{b}, M] \in \Omega - \omega \tag{4a}$$

where

$$\omega = \left\{[0, M] \mid M > 0\right\} \tag{4b}$$

$$\Omega - \omega = \left\{[\boldsymbol{b}, M] \mid M > 0, \boldsymbol{b} \neq 0\right\}. \tag{4c}$$

Then, the generalized likelihood ratio (GLR) test is given by

$$\Lambda(x) = \frac{\max\limits_{\boldsymbol{b}, M \in \Omega - \omega} L(\boldsymbol{b}, M)}{\max\limits_{\boldsymbol{b}, M \in \omega} L(\boldsymbol{b}, M)} \begin{array}{ll} \geq k, & \text{then } H_1 \\ < k, & \text{then } H_0 \end{array} \tag{5}$$

where $k \geq 0$ is the threshold of test.

It is shown, for example, in [8] and [9] that

$$\max_{\boldsymbol{b}, M \in \Omega - \omega} L(\boldsymbol{b}, M) = \frac{1}{2\pi^{NJ/2}|\hat{M}_{\hat{b}}|^{N/2}} \exp\left(\frac{-NJ}{2}\right) \tag{6a}$$

and

$$\max_{\boldsymbol{b}, M \in \omega} L(\boldsymbol{b}, M) = \frac{1}{2\pi^{NJ/2}|\hat{M}_0|^{N/2}} \exp\left(\frac{-NJ}{2}\right) \tag{6b}$$

where

$$\hat{M}_0 = \frac{1}{N}\sum_{n=1}^{N} \boldsymbol{x}(n)\,\boldsymbol{x}^T(n) = \frac{1}{N}XX^T \tag{7a}$$

$$\hat{M}_{\hat{b}} = \frac{1}{N}\sum_{n=1}^{N} \boldsymbol{x}_{\hat{b}}(n)\,\boldsymbol{x}_{\hat{b}}^T(n) = \frac{1}{N}(X - \hat{\boldsymbol{b}}S)(X - \hat{\boldsymbol{b}}S)^T \tag{7b}$$

and

$$\hat{\boldsymbol{b}} = \frac{XS^T}{SS^T} \tag{7c}$$

are the well-known maximum likelihood estimates (MLE's) of the unknown parameters M and $\boldsymbol{b}$ under the hypotheses H_0 and H_1, respectively, e.g., see [9, p. 430, theorem 10.1.1]. Thus a substitution of (6a) and (6b) into (5), yields the GLR test

$$\Lambda(X) = \frac{|\hat{M}_0|^{N/2}}{|\hat{M}_{\hat{b}}|^{N/2}} \begin{array}{ll} \geq k, & \text{then } H_1 \\ < k, & \text{then } H_0. \end{array} \tag{8}$$

Taking the $N/2$th root, this test is evidently equivalent to

$$\lambda(X) = \frac{|\hat{M}_0|}{|\hat{M}_{\hat{b}}} \begin{array}{ll} \geq c, & \text{then } H_1 \\ < c, & \text{then } H_0 \end{array} \tag{9}$$

where $c = k^{2/N}$. A substitution of (7a), (7b), and (7c) into (9) produces the explicit test

$$\lambda(X) = \frac{|XX^T|}{\left|XX^T - \dfrac{(XS^T)(XS^T)^T}{SS^T}\right|} \begin{array}{ll} \geq c, & \text{then } H_1 \\ < c, & \text{then } H_0. \end{array} \tag{10}$$

To further simplify (10) note first that the inverse of XX^T is shown to exist with probability one in the last paragraph of Appendix A. Thus the test ratio in (10) can be considerably simplified by factoring out the determinant of the $J \times J$ matrix XX^T in the denominator to obtain this ratio in the form. This yields

$$\lambda(X) = \frac{|XX^T|}{|XX^T|\left|I - (XX^T)^{-1/2}\dfrac{(XS^T)(XS^T)^T}{SS^T}(XX^T)^{-1/2}\right|} = \frac{1}{1 - \dfrac{(XS^T)^T(XX^T)^{-1}(XS^T)}{SS^T}}. \tag{11}$$

The last equation follows from a well-known determinant identity, also see [3, appendix A]. Clearly the test in (11) is equivalent finally to the test

$$r(X) = \frac{(XS^T)^T(XX^T)^{-1}(XS^T)}{SS^T} \begin{array}{ll} \geq r_0, & \text{then } H_1 \\ < r_0, & \text{then } H_0. \end{array} \tag{12}$$

For $J = 1$ this reduces a well-known CFAR test for one receiving channel.

IV. Detection and False Alarm Probabilities of Test

In order to find the probability density function of the test r in (12) on both hypotheses H_0 and H_1, one assumes the $\boldsymbol{b}$ and M are known and start by noting that

$$\text{cov}\,[\boldsymbol{x}(n)|H_i] = M \qquad \text{for } i = 0, 1. \tag{13}$$

Also

$$\boldsymbol{E}[\boldsymbol{x}(n)|H_0] = \boldsymbol{E}[\boldsymbol{x}^0(n)] = 0 \tag{14a}$$

and

$$\boldsymbol{E}[\boldsymbol{x}(n)|H_1] = \boldsymbol{E}[\boldsymbol{x}^0(n) + \boldsymbol{b}s(n)] = \boldsymbol{b}s(n) \tag{14b}$$

or in terms of X

$$\boldsymbol{E}[X|H_0] = 0 \quad \text{and} \quad \boldsymbol{E}[X|H_1] = \boldsymbol{b}S. \tag{15}$$

Next perform a whitening procedure on $\boldsymbol{x}(n)$ by defining

$$z(n) = M^{-1/2}\boldsymbol{x}(n), \qquad \text{for } n = 1, 2, \cdots, N \tag{16a}$$

i.e., let

$$Z = [z(1)\, z(2), \cdots, z(N)] = M^{-1/2}X. \tag{16b}$$

The whitening procedure in (16a) and the assumption that the residual clutter samples in the spatial coordinates are mutually independent produce the result

$$\text{cov}\,[z_i(m)\, z_j(n)] = \boldsymbol{\delta}(i, j)\, \boldsymbol{\delta}(n, m) \tag{17}$$

for $i, j = 1, 2, \cdots, J$ and $m, n = 1, 2, \cdots, N$. Here $\boldsymbol{\delta}(n, m)$ is the Kronecker delta function defined by

$$\boldsymbol{\delta}(n, m) = \begin{cases} 1 & \text{if } n = m \\ 0 & \text{otherwise} \end{cases} \tag{18}$$

and $z_i(m)$ is the ith element of vector $z(m)$. Then by (15) to (18)

$$\boldsymbol{E}[Z|H_0] = 0 \tag{19}$$

$$\boldsymbol{E}[Z|H_1] = M^{-1/2}\boldsymbol{b}S \tag{20}$$

and

$$\text{cov}\,[z(n)|H_i] = I_J \qquad \text{for } i = 0, 1 \tag{21}$$

where I_J is the $J \times J$ identity matrix.

Evidently by the transformation in (16b) the test function in (12) becomes

$$r = \frac{(ZS^T)^T(ZZ^T)^{-1}(ZS^T)}{(SS^T)} \begin{array}{ll} \geq r_0, & \text{then } H_1 \\ < r_0, & \text{then } H_0. \end{array} \tag{22}$$

Since SS^T is a positive scalar, at this point it simplifies matters to normalize the signal vector S by letting

$$S_1 = \frac{S}{(SS^T)^{1/2}}. \tag{23}$$

Then the test function in (22) becomes, using (23)

$$r = (ZS_1^T)^T(ZZ^T)^{-1}(ZS_1^T). \tag{24}$$

By (23) the sum-of-squares norm of S_1 is given by $\|S_1\| = 1$. Hence, S_1 is a unit row vector in the "direction" of vector S.

Now consider the $N \times N$ orthonormal matrix

$$U = \begin{bmatrix} S_1 \\ Q \end{bmatrix} \tag{25}$$

where Q is a $(N - 1) \times N$ matrix, composed of some set of orthonormal row vectors, and such that

$$S_1 Q^T = 0.$$

Hence the matrix U carries out rotations in N-dimensional space, in such a manner that unit vector S_1 is transformed into the new unit vector,

$$S_1 U^T = [1, 0, \cdots, 0]. \tag{26}$$

Now apply transformation U to Z by letting

$$V = ZU^T = [\boldsymbol{v}(1), \boldsymbol{v}(2), \cdots, \boldsymbol{v}(N)]. \tag{27}$$

Then the test function r in (24) reduces to

$$r = \boldsymbol{v}(1)^T(VV^T)^{-1}\, \boldsymbol{v}(1). \tag{28}$$

For more details on how U acts upon signal S_1 and X or Z under H_1, see Appendix A.

The covariance matrix of $\boldsymbol{v}(n)$, for $n = 1, 2, \cdots, N$ is similar to that of $z(n)$, the only change of the statistics of the $\boldsymbol{v}(n)$ from that of the $z(n)$ is their mean values under hypothesis H_1. This mean is derived from

$$\begin{aligned} \boldsymbol{E}[V|H_1] &= \boldsymbol{E}[ZU^T|H_1] \\ &= M^{-1/2}\boldsymbol{b}S_1U^T(SS^T)^{1/2} \\ &= M^{-1/2}\boldsymbol{b}[1, 0, \cdots, 0](SS^T)^{1/2} \\ &= [M^{-1/2}\boldsymbol{b}(SS^T)^{1/2}, \boldsymbol{0}, \cdots, \boldsymbol{0}]. \end{aligned} \tag{29}$$

From (17) and (29) a figure of merit or what might be termed the generalized signal-to-noise ratio (GSNR) of the test, is developed as follows:

$$(\mathrm{GSNR}) = E[\boldsymbol{v}^T(1)|H_1]\, E[\boldsymbol{v}(1)|H_1] = (\boldsymbol{b}^T M^{-1}\boldsymbol{b})\|S\|^2 \triangleq a. \tag{30}$$

Now consider a further simplification of the test function r in (28). First separate matrix V into two parts, $V = [\boldsymbol{v}(1), D]$ and $D = [\boldsymbol{v}(2), \cdots, \boldsymbol{v}(N)]$, in such a manner that

$$VV^T = \boldsymbol{v}(1)\,\boldsymbol{v}^T(1) + \sum_{n=2}^{N} \boldsymbol{v}(n)\,\boldsymbol{v}^T(n) = \boldsymbol{v}(1)\,\boldsymbol{v}^T(1) + \boldsymbol{\Phi}. \tag{31a}$$

Here

$$\boldsymbol{\Phi} = DD^T = \sum_{n=2}^{N} \boldsymbol{v}(n)\,\boldsymbol{v}^T(n) \tag{31b}$$

is a nonsingular $J \times J$ matrix, since $N - 1 \geq J$ and $\boldsymbol{\Phi}$ is obviously Wishart distributed (see Appendix A for more, but similar, details).

A well-known matrix inversion identity applied to (31a) produces the result

$$(VV^T)^{-1} = [\boldsymbol{v}(1)\,\boldsymbol{v}^T(1) + \boldsymbol{\Phi}]^{-1} = \left[I - \frac{\boldsymbol{\Phi}^{-1}\boldsymbol{v}(1)\,\boldsymbol{v}^T(1)}{1 + \boldsymbol{v}^T(1)\,\boldsymbol{\Phi}^{-1}\boldsymbol{v}(1)}\right]\boldsymbol{\Phi}^{-1}. \tag{32}$$

A substitution of (32) into the test function r in (28) yields the test function r as the new expression

$$r = \frac{\boldsymbol{v}^T(1)\,\boldsymbol{\Phi}^{-1}\boldsymbol{v}(1)}{1 + \boldsymbol{v}^T(1)\,\boldsymbol{\Phi}^{-1}\boldsymbol{v}(1)} = \frac{r_1}{1 + r_1} \tag{33}$$

where

$$r_1 = \boldsymbol{v}^T(1)\,\boldsymbol{\Phi}^{-1}\boldsymbol{v}(1). \tag{34}$$

It is desired now to find the probability density, $f(r_1|H_1)$, of r_1 in (34). First reexpress (34) in the form

$$r_1 = \|\boldsymbol{v}(1)\|^2\left(\frac{\boldsymbol{v}^T(1)}{\|\boldsymbol{v}(1)\|}(DD^T)^{-1}\frac{\boldsymbol{v}(1)}{\|\boldsymbol{v}(1)\|}\right). \tag{35}$$

Then normalize the J-component vector $\boldsymbol{v}(1)$ as follows:

$$\xi = \frac{\boldsymbol{v}(1)}{\|\boldsymbol{v}(1)\|}. \tag{36}$$

Hence by (36) one obtains r_1 in (35) in the form

$$r_1 = \|\boldsymbol{v}(1)\|^2(\xi^T(DD^T)^{-1}\xi) = \|\boldsymbol{v}(1)\|^2 e \tag{37a}$$

where

$$e = (\xi^T(DD^T)^{-1}\xi). \tag{37b}$$

Now one can further process the term e in (37b) by conditioning on the elements of $\boldsymbol{v}(1)$ so that ξ can be treated as a normalized constant vector. Then since ξ has unity magnitude, there exists a $J \times J$ orthonormal matrix U_1 such that (see also [8], [9])

$$U_1\xi = [1, 0, \cdots, 0]^T. \tag{38}$$

Next apply this transformation to matrix D, defined before (31a), by letting

$$H = U_1 D = U_1[\boldsymbol{v}(2), \cdots, \boldsymbol{v}(N)]. \tag{39}$$

Then the term e in (37b) has the simple form,

$$e = \xi^T(DD^T)^{-1}\xi = [1, 0, \cdots, 0](HH^T)^{-1}[1, 0, \cdots, 0]^T. \tag{40}$$

Clearly H in (39) has exactly the same statistical properties as D under the assumption that $\boldsymbol{v}(1)$ is given.

Now partition H as follows:

$$H = \begin{bmatrix} \boldsymbol{h}_A^T \\ H_B \end{bmatrix} \tag{41a}$$

where $\boldsymbol{h}_A$ is the $(N-1)$-column vector and H_B is the $(J-1) \times (N-1)$ matrix. Then

$$(HH^T)^{-1} = \begin{bmatrix} \boldsymbol{h}_A^T\boldsymbol{h}_A & \boldsymbol{h}_A^T H_B^T \\ H_B\boldsymbol{h}_A & H_B H_B^T \end{bmatrix}^{-1} \triangleq \begin{bmatrix} R_{AA} & R_{AB} \\ R_{BA} & R_{BB} \end{bmatrix}. \tag{41b}$$

According to the Frobenius relations, e.g., see [8] or [9], for a partitioned matrix

$$R_{AA} = [\boldsymbol{h}_A^T\boldsymbol{h}_A - \boldsymbol{h}_A^T H_B^T(H_B H_B^T)^{-1}H_B\boldsymbol{h}_A]^{-1} = [\boldsymbol{h}_A^T(I - H_B^T(H_B H_B^T)^{-1}H_B)\boldsymbol{h}_A]^{-1} = \frac{1}{\boldsymbol{h}_A^T(I - H_B^T(H_B H_B^T)^{-1}H_B)\boldsymbol{h}_A} = \frac{1}{\boldsymbol{h}_A^T P_1 \boldsymbol{h}_A}. \tag{42}$$

A substitution of (41b) and (42) into (40) yields

$$e = \frac{1}{\boldsymbol{h}_A^T P_1 \boldsymbol{h}_A} \tag{43}$$

where $P_1 \triangleq I - H_B^T(H_B H_B^T)^{-1}H_B$ is a projection operator such that $P_1^2 = P_1$ and $\mathrm{Tr}\,(P_1) = N - J$. In the same manner used for the projection matrix P in [1] it is not difficult to show that P_1 has $N - J$ unity eigenvalues and $J - 1$ zero eigenvalues. Thus P_1 can be diagonalized to the form

$$U_2^T P_1 U_2 = \Lambda_1 = \begin{bmatrix} I_{N-J} & 0 \\ 0 & 0_{J-1} \end{bmatrix}. \tag{44}$$

By arguments similar to those used previously in [1], one finds also, under the assumption that $\boldsymbol{v}(1)$ and P_1 are given, that the random variable

$$\frac{1}{e} = \boldsymbol{h}_A^T P_1 \boldsymbol{h}_A = \boldsymbol{\eta}^T\boldsymbol{\eta} = \sum_{i=1}^{N-J} \eta_i^2 \tag{45a}$$

where e is defined in (43), and

$$\boldsymbol{\eta} \triangleq \Lambda_1^{1/2} U_2^T \boldsymbol{h}_A \tag{45b}$$

is a $(N-1)$-column vector with the last $(J-1)$ components equal to zero. The conditional joint probability density function of the first $(N-J)$ nonzero elements of $\boldsymbol{\eta}$ is subject to the normal density function, $N(\mathbf{0}, I_{N-J})$, i.e.,

$$p_{\boldsymbol{\eta}}(\eta_1, \cdots, \eta_{(N-J)} | \boldsymbol{v}(1), P_1) = N(\mathbf{0}, I_{N-J}). \quad (46)$$

But also $p_{\boldsymbol{\eta}}(\eta_1, \cdots, \eta_{(N-J)} | \boldsymbol{v}(1), P_1)$ in (46) does not depend on $\boldsymbol{v}(1)$ and P_1, so that $\boldsymbol{\eta}$ and $1/e$ in (45a) must be statistically independent of the vector $\boldsymbol{v}(1)$ and matrix P_1. Hence $1/e$ in (45a) is chi-squared distributed.

Also by (37a) and (43) one obtains the ratio r_1 in the form

$$r_1 = \frac{\boldsymbol{v}^T(1)\,\boldsymbol{v}(1)}{\boldsymbol{\eta}^T\boldsymbol{\eta}} = \frac{\sum_{j=1}^{J} v_j^2(1)}{\sum_{i=1}^{N-J} \eta_i^2} \quad (47)$$

in terms of the magnitudes of these two vectors only. Thus by the independence of vectors $\boldsymbol{\eta}$ and $\boldsymbol{v}(1)$ and a use of [14, p. 52, corollary 2], one has the probability density function

$$f(r_1|H_1) = \frac{r_1^{(J-2/2)} e^{(-a/2)}}{B\left(\frac{N-J}{2}, \frac{1}{2}J\right)(1+r_1)^{N/2}} \cdot {}_1F_1\left(\frac{N}{2}; \frac{1}{2}J; \frac{ar_1}{2(1+r_1)}\right) \quad (48)$$

of r_1 in (34) under hypothesis H_1 where ${}_1F_1(a; b; x)$ is the confluent hypergeometric function. This is a noncentral F distribution. The value a in [14, p. 52] is given by $a = \| \boldsymbol{E}\{\boldsymbol{v}(1) | H_1\} \|^2$. This agrees with the definition of the GSNR, a, given in (30).

Finally by using the relationship of r_1 to r in (33) the probability density function of the test function r under hypothesis H_1 is given by

$$f(r|H_1) = \frac{\Gamma\left(\frac{N}{2}\right)}{\Gamma\left(\frac{N-J}{2}\right)\Gamma\left(\frac{J}{2}\right)} \cdot (1-r)^{(N-J-2/2)} r^{(J-2/2)} e^{(-a/2)} \cdot {}_1F_1\left(\frac{N}{2}; \frac{J}{2}; \frac{ar}{2}\right) \quad (49)$$

for $0 < r < 1$ where a is the generalized SNR in (30). This says that r is subject to a noncentral beta-distribution. Clearly, if no signal is present, then $a = 0$. Thus (49) reduces in the H_0 hypothesis to a standard beta-function density of form

$$f(r|H_0) = \frac{\Gamma\left(\frac{N}{2}\right)}{\Gamma\left(\frac{N-J}{2}\right)\Gamma\left(\frac{J}{2}\right)} (1-r)^{(N-J-2/2)} r^{(J-2/2)} \quad \text{for } 0 < r < 1. \quad (50)$$

Finally, in terms of the above probability density functions in (49) and (50) the probability of a false alarm is found by

$$P_{\mathrm{FA}} = \int_{r_0}^{1} f(r|H_0)\, dr \quad (51)$$

and the probability of detection by

$$P_D = \int_{r_0}^{1} f(r|H_1)\, dr. \quad (52)$$

V. Performance Analysis

Performance curves of the probability of detection in (52) versus SNR for a given false alarm probability with respect to various values of N and J are computed in this section. First, in Fig. 4 the probability of detection is calculated as a function of the generalized signal-to-noise ratio (GSNR) a in (30) for several different values of parameter N. These curves demonstrate that, for a fixed GSNR, a, the CFAR detector has a higher detection probability if more samples are used.

In Appendix B the performance improvement in the GSNR obtained from a full data array X is compared with that of partial data array of less rows in a theorem. This theorem shows that the gain in (30) of the test in (12) using the full data array X is always greater than the GSNR of the same test which uses a data array X_q composed of any subset of the rows of data array X. With this theorem the probability of detection in a single scene of maximum SNR can be compared with the probability of detection in J correlated scenes. To accomplish this, the GSNR in J correlated scenes is related to the maximum SNR of the J scenes in a manner shown next.

Let the maximum signal-to-noise ratio in the J correlated scenes be

$$a' = \frac{b_1^2 \|S\|^2}{\sigma_1^2}. \quad (53)$$

By (30) the GSNR for J bands is given by

$$a_J = [b_1 \boldsymbol{b}_2^T] M^{-1} [b_1 \boldsymbol{b}_2^T]^T \|S\|^2 \quad (54)$$

where

$$M = \boldsymbol{E}\left\{\begin{bmatrix} x_1(n) - \boldsymbol{E}x_1(n) \\ \boldsymbol{x}(n) - \boldsymbol{E}\boldsymbol{x}(n) \end{bmatrix}\begin{bmatrix} x_1(n) - \boldsymbol{E}x_1(n) \\ \boldsymbol{x}(n) - \boldsymbol{E}\boldsymbol{x}(n) \end{bmatrix}^T\right\} = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix}$$

$$\boldsymbol{x}(n) \triangleq [x_2(n)\, x_3(n), \cdots, x_J(n)]^T,$$

$$\boldsymbol{b}_2 \triangleq [b_2\, b_3, \cdots, b_J]^T$$

and a_K denotes the GSNR using K scenes. Next define a general normalized correlation coefficient vector by

$$\boldsymbol{\rho} \triangleq K_{11}^{-1/2} K_{22}^{-1/2} K_{21} \quad (55a)$$

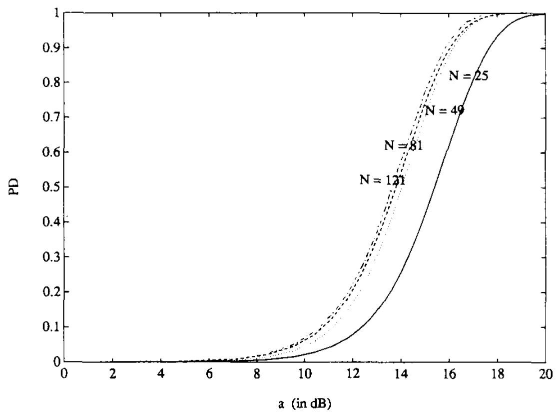


Fig. 4. Probability of detection versus GSNR. $J = 2$ and $P_{FA} = 10^{-5}$.

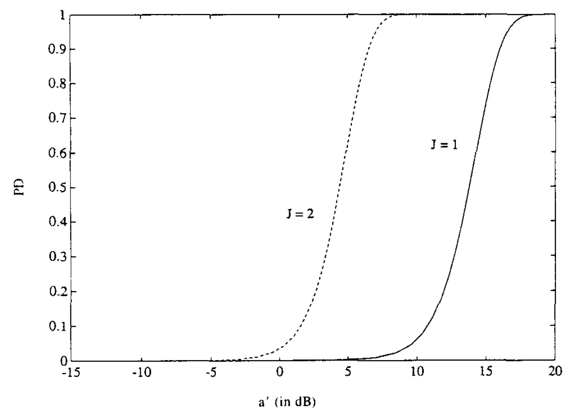


Fig. 5. Probability of detection versus SNR. $\rho = 0.95$, $\lambda = 0.2$, and $P_{FA} = 10^{-5}$.

and define a $(J - 1)$-vector λ as

$$\lambda = \frac{K_{11}^{1/2}}{b_1} K_{22}^{-1/2} b_2. \tag{55b}$$

In terms of vector λ in (55b), the ratio of the GSNR in the other $J - 1$ scenes to the SNR in the primary scene of maximum SNR is given by

$$\frac{a_{J-1}}{a'} = \lambda^T \lambda = \frac{K_{11}}{b_1^2} b_2^T K_{22}^{-1} b_2. \tag{56}$$

Finally, the generalized gain G, given by (B10) in the Theorem of Appendix B for $q = 1$, can be reexpressed as

$$G = \frac{1 - 2\lambda^T\rho + \lambda^T\lambda(1 - \rho^T\rho) + (\lambda^T\rho)^2}{1 - \rho^T\rho}. \tag{57}$$

Note by the definition of ρ in (55a) that $\rho^T\rho < 1$ so that G in (57) is well defined.

Consider now the case $J = 2$. For this

$$a = [b_1\ b_2] M_2^{-1} [b_1\ b_2]^T \|S\|^2 \tag{58}$$

where

$$M_2 = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix} \quad \text{and}$$

$$M_2^{-1} = \frac{1}{\sigma_1^2\sigma_2^2 - \sigma_{12}^2} \begin{bmatrix} \sigma_2^2 & -\sigma_{12} \\ -\sigma_{21} & \sigma_1^2 \end{bmatrix}. \tag{59}$$

As a consequence one has

$$\begin{aligned} a &= [b_1\ b_2] M_2^{-1} [b_1\ b_2]^T \|S\|^2 \\ &= \frac{(b_1^2\sigma_2^2 + b_2^2\sigma_1^2 - 2b_1b_2\sigma_{12})\|S\|^2}{\sigma_1^2\sigma_2^2 - \sigma_{12}^2} \\ &= \frac{a'}{1 - \dfrac{\sigma_{12}^2}{\sigma_1^2\sigma_2^2}} \left(1 + \frac{b_2^2}{b_1^2}\frac{\sigma_1^2}{\sigma_2^2} - 2\frac{b_2}{b_1}\frac{\sigma_{12}}{\sigma_2^2}\right). \end{aligned} \tag{60}$$

In terms of the normalized correlation coefficient

$$\rho = \frac{\sigma_{12}}{\sigma_1\sigma_2} \tag{61}$$

and the ratio

$$\lambda^2 = \frac{b_2^2\|S\|^2/\sigma_2^2}{b_1^2\|S\|^2/\sigma_1^2} \tag{62}$$

of SNR in the second scene to SNR in the primary scene of maximum SNR, (57) can be reduced as

$$G = \frac{(1 + \lambda^2 - 2\lambda\rho)}{1 - \rho^2} = 1 + \frac{(\lambda - \rho)^2}{1 - \rho^2}. \tag{63}$$

It means that the gain in signal-to-noise ratio of a detector, which uses $J = 2$ scenes, over a single detector which uses that scene with the maximum SNR. Fig. 5 illustrates the probability of detection for a false alarm probability of $P_{FA} = 10^{-5}$ as a function of a' for $J = 1, 2$, $N = 49$, $\lambda = 0.2$, and $\rho = 0.95$, i.e., for $G = 6.77$. This shows that this detector using 2 scenes with $(\text{SNR})_2 = \frac{1}{5}(\text{SNR})_1$ has an approximate 8.5 dB SNR improvement over a detector which uses the single scene with maximum SNR.

A comparison of the GSNR gain in (63) for $J = 2$ is now made with [1, eq. 27] for $K = 1$ single noise-only reference scene case in [1]. By [1, eq. 27] the SNR for a scene with signal, using one reference without signal, is given by

$$a = (\sigma_y^2/\sigma_{y|x}^2)a' = \frac{a'}{(1 - \rho^2)} \triangleq \tilde{G}a' \tag{64}$$

in terms of ρ and a', the SNR in the scene with signal. By (63) the GSNR gain depends on ρ, the correlation coefficient, but also on λ, the ratio of SNR's in both scenes. It is evident from (63) and (64), that as long as the inequality $\lambda/2 < \rho < 1$ holds, the gain in SNR in (63) is always less than or equal to the gain in the SNR given in (64).

Using the correlation coefficient $\rho = 0.81$, a computational comparison is made in Fig. 6 with the results given

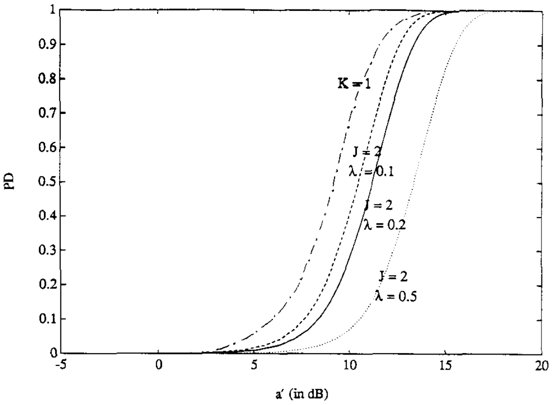


Fig. 6. Probability of detection versus SNR. $\rho = 0.81$ and $P_{FA} = 10^{-5}$.

in [1] for $K = 1$. These curves in Fig. 6 illustrate that for the same ρ the probability of detection curves for the new CFAR detector are limited to the probability of detection curves given in [1] for $K = 1$ if $\lambda/2 < \rho < 1$ holds. The leftmost curve in Fig. 6 is derived under the assumption that there is no signal in the reference scene.

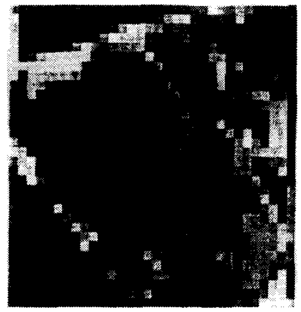
Figs. 7(a) and (b) show typical 32×32 subimages in two different optical bands (the green and red bands) of the San Diego area. A 5×5 signal with the pattern, given in Fig. 8, is implanted in both of these green and red images with $(\text{SNR})_1 = 0.26$ and $(\text{SNR})_2 = \frac{1}{2}(\text{SNR})_1 = 0.5 \times 0.26 = 0.13$. A local mean is subtracted from both of these subimages. The resulting residual images are approximately zero-mean Gaussian processes. The CFAR test given in (12) is calculated for each pixel. The test statistic of this CFAR test is plotted pixel by pixel in Fig. 7(c). A target is detected with a threshold determined by $P_{FA} = 10^{-5}$. The results of this test are illustrated in Fig. 7(d) which shows that the target was, in fact, detected.

In order to demonstrate the theoretical GSNR improvement in (63), a computer simulation was performed to determine the required SNR in the primary scene of maximum SNR needed to detect a target in single primary scene and in $J = 2$ scenes. A computer simulation was made similar to that developed in [1]. The results of this new CFAR test are shown in Table I, where $\lambda = 0.2$ and $\lambda = 0.5$, respectively. Using 5 different subimages the average of the resulting five signal-to-noise ratios of a detector using $J = 2$ scenes over the single primary scene is 5.19 dB for $\lambda = 0.2$ and 3.25 dB for $\lambda = 0.5$.

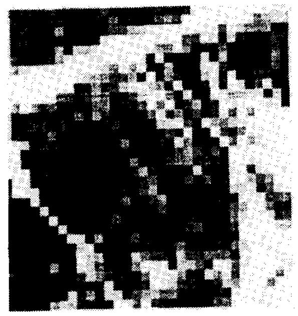
VI. Conclusions

Under the same assumptions for optical noise clutter used in [1], a CFAR algorithm is developed for detecting the presence of a dim optical signal of nonzero intensity in J signal-plus-noise bands or channels. This new algorithm is more flexible and more practical than the one given in [1]. If $J = 1$, the resulting test reduces to the standard normalized matched filter test for finding a signal in clutter of unknown and varying intensity.

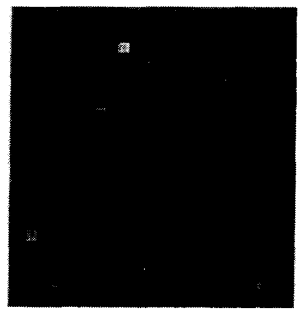
Both theoretical and computer simulation results show that the SNR improvement factor of this new algorithm



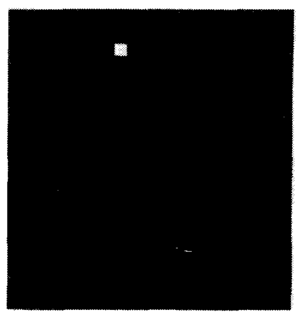
(a)



(b)



(c)



(d)

Fig. 7. San Diego color images. (a) Green subimage with target. (b) Red subimage with target. (c) CFAR test values of each pixel. (d) Target detected by the threshold.

$$\begin{bmatrix} 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \end{bmatrix}$$

Fig. 8. Target signal template.

TABLE I
CFAR ALGORITHM SIMULATION RESULTS FOR THE OPTICAL COLOR IMAGE WITH $N = 49$ SAMPLES FOR COVARIANCE ESTIMATION

Subimage Location	ρ	J = 1 (dB)	J = 2 λ = 0.2 (dB)	J = 2 λ = 0.5 (dB)
(100, 10)	0.94	6.67	0.96	4.70
(234, 30)	0.92	9.34	2.04	5.91
(50, 310)	0.97	7.32	5.18	3.98
(200, 200)	0.95	2.30	-4.32	-5.91
(65, 1)	0.93	3.80	-4.20	0.72
Average	0.94	6.58	1.39	3.33
Improvement Factor		—	5.19	3.25

using multiple band scenes over the single scene of maximum SNR can be substantial. The comparison of SNR gain between this new detection algorithm and the one given in [1] illustrates that for the same correlation coefficient of related scenes ρ the probability of detection curves for the new CFAR detector are limited to the probability of detection curves given in [1] if $\lambda/2 < \rho < 1$ holds.

Appendix A
Properties of Orthogonal Matrix U

To see explicitly how orthonormal transformation $U = \left[\begin{smallmatrix} S_1 \\ Q \end{smallmatrix}\right]$ in (25) acts upon the normalized version S_1 of signal S in (1b) and data matrix X under H_1, the signal-plus-noise situation, let

$$X \triangleq \boldsymbol{b}_1 S_1 + X_0 \tag{A1}$$

where $\boldsymbol{b}_1 = \boldsymbol{b}(SS^T)^{1/2}$, $S_1 = S(SS^T)^{-1/2}$, and X_0 is data matrix X when $\boldsymbol{b}_1 = 0$ or hypothesis H_0 is true. Since

$$S_1 U^T = S_1[S_1^T, Q^T] = [S_1 S_1^T, S_1 Q^T] = [1, 0, \cdots, 0], \tag{A2}$$

then a multiplication of X in (A1) on the right by U^T yields

$$\begin{aligned} Y = XU^T &= [\boldsymbol{b}_1 S_1 + X_0][S_1^T, Q^T] \\ &= [(\boldsymbol{b}_1 S_1 + X_0)S_1^T, (\boldsymbol{b}_1 S_1 + X_0)Q^T] \\ &= [\boldsymbol{b}_1 + X_0 S_1^T, X_0 Q^T]. \end{aligned} \tag{A3}$$

Evidently the action of U on X is to send signal $\boldsymbol{b}_1 S_1$ to $\boldsymbol{b}_1[1, 0, \cdots, 0]$ plus a noise term $X_0 S_1^T$ into the first column only of the transformed data matrix Y. The remaining $N - 1$ columns of Y, namely $X_0 Q^T$, constitute a signal-free $J \times (N - 1)$ matrix from which, as it is shown next, an estimate of the covariance matrix M can be found.

First note by the definition of Y and U that

$$\begin{aligned} YY^T &= X[S_1^T, Q^T][S_1^T, Q^T]^T X^T \\ &= (XS_1^T)(XS_1^T)^T + (XQ^T)(XQ^T)^T. \end{aligned} \tag{A4.1}$$

Finally under H_1 by (A3) one has

$$\begin{aligned} YY^T &= [(\boldsymbol{b}_1 S_1 + X_0)S_1^T, X_0 Q^T][(\boldsymbol{b}_1 S_1 + X_0)S_1^T, X_0 Q^T]^T \\ &= [(\boldsymbol{b}_1 S_1 + X_0)S_1^T][(\boldsymbol{b}_1 S_1 + X_0)S_1^T]^T \\ &\quad + (X_0 Q^T)(X_0 Q^T)^T \\ &= (XS_1^T)(XS_1^T)^T + (X_0 Q^T)(X_0 Q^T)^T. \end{aligned} \tag{A4.2}$$

The above identities (A4.1) to (A4.2) evidently establish the following relations:

$$\begin{aligned} (XQ^T)(XQ^T)^T &= (X_0 Q^T)(X_0 Q^T)^T \\ &= XX^T - (XS_1^T)(XS_1^T)^T \triangleq G. \end{aligned} \tag{A5}$$

Taking the expected value of the left side of (A5) yields, by the use of the independence of the columns of X

$$\boldsymbol{E}(XQ)(XQ)^T = \boldsymbol{E}(X_0 Q)(X_0 Q)^T = (N - 1)M \tag{A6}$$

where M is $J \times J$ covariance matrix of the discrete vector process $\boldsymbol{x}(n)$. This result shows that

$$\hat{M}_Q = \frac{1}{N - 1}(XQ)(XQ)^T \tag{A7}$$

is an unbiased estimate of M with $(N - 1)$ degrees of freedom under both of the noise-only and signal-plus-noise hypotheses H_0 and H_1, respectively. Note finally by [15, theorem 5.17] if $N - 1 \geq J$ or $N > J$ that the $K = (1/2)J(J + 1)$, distinct elements of $\hat{M}_Q$ are jointly Wishart distributed over the K-dimensional space of positive definite matrices (see also [3, appendix]). Hence $\hat{M}_Q$ in (A7) is positive definite with probability one.

The inverse of XX^T can be shown explicitly to likewise exist by noting the following:

$$\begin{aligned} XX^T &= G + (XS_1^T)(XS_1^T)^T \\ &= (I + (XS_1^T)(XS_1^T)^T G^{-1})G \triangleq BG. \end{aligned}$$

Evidently XX^T is invertible since clearly both factors B^{-1} and G^{-1} exist with probability one.

Appendix B
Analysis of GSNR Gain

The gain in the GSNR obtained from a full data array X of J rows compared with that of a partial data array of less rows is considered in the next theorem.

Theorem: The GSNR in (30) of the test in (12) using the full data array X is greater than the GSNR of the same test which uses a data array X_q composed of any subset of the rows of the data array X.

Proof: Let $\{i_1, i_2, \cdots, i_q\}$ is the subset of row indices of data array X corresponding to the rows of some partial data array X_q. Next let the GSNR of the test using a subset X_q of data array rows of X be a_q. Then

$$\begin{aligned} a_q &= [b_{i_1}, b_{i_2}, \cdots, b_{i_q}]M_q^{-1}[b_{i_1}, b_{i_2}, \cdots, b_{i_q}]^T \\ &\triangleq \boldsymbol{b}_1^T K_{11}^{-1} \boldsymbol{b}_1 \end{aligned} \tag{B1}$$

where

$$b_1 \triangleq [b_{i_1}, b_{i_2}, \cdots, b_{i_q}]^T \tag{B2.1}$$

$$K_{11} \triangleq M_q = E[x_q(n) - Ex_q(n)][x_q(n) - Ex_q(n)]^T \tag{B2.2}$$

and

$$x_q(n) = [x_{i_1}(n), x_{i_2}(n), \cdots, x_{i_q}(n)]^T. \tag{B2.3}$$

By the use of a permutation matrix π one can always express the GSNR of the test in (30), using the full data array, as follows:

$$\begin{aligned} a_J &= b^T M^{-1} b = [\pi^{-1}(\pi b)]^T M^{-1} [\pi^{-1}(\pi b)] \\ &= (\pi^{-1}\tilde{b})^T M^{-1} (\pi^{-1}\tilde{b}) \\ &= \tilde{b}^T K^{-1} \tilde{b} = [b_1^T\ b_2^T] \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix}^{-1} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \end{aligned} \tag{B3}$$

where

$$K = \pi M \pi^T \quad \text{and} \quad \tilde{b} = \pi b \triangleq \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

Now define the following orthogonal projection operators

$$P = \begin{bmatrix} 0 & 0 \\ -K_{21}K_{11}^{-1} & I \end{bmatrix} \tag{B4.1}$$

and

$$Q \triangleq I - P = \begin{bmatrix} I & 0 \\ K_{21}K_{11}^{-1} & 0 \end{bmatrix} \tag{B4.2}$$

on the inner product space R^J where the inner product is $(u, v) \triangleq u^T K^{-1} v$. Then one can rewrite $\tilde{b}$ in the form

$$\tilde{b} = P\tilde{b} + (I - P)\tilde{b} = P\tilde{b} + Q\tilde{b} \tag{B5}$$

where

$$P\tilde{b} = \begin{bmatrix} 0 \\ b_2 - K_{21}K_{11}^{-1}b_1 \end{bmatrix} \tag{B6.1}$$

and

$$Q\tilde{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} - \begin{bmatrix} 0 \\ b_2 - K_{21}K_{11}^{-1}b_1 \end{bmatrix} = \begin{bmatrix} b_1 \\ K_{21}K_{11}^{-1}b_1 \end{bmatrix}. \tag{B6.2}$$

By (B6.1) and (B6.2), it is not difficult to verify that

$$(P\tilde{b})^T K^{-1} (Q\tilde{b}) \equiv 0 \tag{B7}$$

so that by (B3)

$$a_J = \tilde{b}^T K^{-1} \tilde{b} = (P\tilde{b})^T K^{-1} (P\tilde{b}) + (Q\tilde{b})^T K^{-1} (Q\tilde{b}). \tag{B8}$$

Next by using the Frobenius relations, e.g., see [8] or [9], one can show that

$$(P\tilde{b})^T K^{-1} (P\tilde{b}) = (b_2 - K_{21}K_{11}^{-1}b_1)^T \cdot K^{22}(b_2 - K_{21}K_{11}^{-1}b_1) \tag{B9.1}$$

$$(Q\tilde{b})^T K^{-1} (Q\tilde{b}) = b_1^T K_{11}^{-1} b_1 \tag{B9.2}$$

where

$$K^{22} = (K_{22} - K_{21}K_{11}^{-1}K_{12})^{-1}.$$

Thus a substitution of (B9.1) and (B9.2) into (B8) yields finally

$$\begin{aligned} a_J &= (b_2 - K_{21}K_{11}^{-1}b_1)^T K^{22} (b_2 - K_{21}K_{11}^{-1}b_1) + b_1^T K_{11}^{-1} b_1 \\ &= (b_1^T K_{11}^{-1} b_1) \\ &\quad \cdot \left[\frac{(b_2 - K_{21}K_{11}^{-1}b_1)^T K^{22} (b_2 - K_{21}K_{11}^{-1}b_1)}{b_1^T K_{11}^{-1} b_1} + 1 \right] \\ &= a_q G \end{aligned}$$

where

$$G \triangleq \frac{(b_2 - K_{21}K_{11}^{-1}b_1)^T K^{22} (b_2 - K_{21}K_{11}^{-1}b_1)}{b_1^T K_{11}^{-1} b_1} + 1 \geq 1. \tag{B10}$$

Thus the theorem is proved. The method used here to prove the theorem is similar to the method used to factor a multivariate Gaussian density distribution function into the product of a conditional and a unconditional probability density function.

Acknowledgment

The authors wish to thank Dr. E. J. Kelly, Lincoln Laboratory, Massachusetts Institute of Technology, for several valuable discussions. They also wish to acknowledge the reviewers for important suggested improvements made in this paper during the review process.

References

[1] J. Y. Chen and I. S. Reed, "A detection algorithm for optical targets in clutter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-23, no. 1, Jan. 1987.

[2] J. Y. Chen, "Adaptive detection of optical and radar targets using reference processes," Ph.D. dissertation, Univ. Southern California, Los Angeles, CA, Apr. 1986.

[3] A. Margalit, I. S. Reed, and R. M. Gagliardi, "Adaptive optical target detection using correlated images," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-21, May 1985.

[4] A. Margalit, "Adaptive detection of stationary optical and IR targets using correlated scenes," Ph.D. dissertation, Univ. Southern California, Los Angeles, CA, Nov. 1984.

[5] B. R. Hunt and T. M. Cannon, "Nonstationary assumptions of Gaussian models of images," *IEEE Trans. Syst., Man, Cybern.*, Dec. 1976.

[6] K. M. Yang, "Problems in detection and estimation," Ph.D. dissertation, Univ. Southern California, Los Angeles, CA, Apr. 1987.

[7] L. E. Brennan and I. S. Reed, "An adaptive array signal algorithm for communications," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-18, Jan. 1982.

[8] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, 2nd ed. New York: Wiley, 1985.

[9] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*. New York: Wiley, 1982.

[10] E. J. Kelly, "An adaptive detection algorithm," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-22, Mar. 1986.

[11] E. J. Kelly and K. M. Forsythe, "Adaptive detection and parameter estimation for multidimensional signal models," Lincoln Lab., Tech. Rep. 848, Apr. 1989.

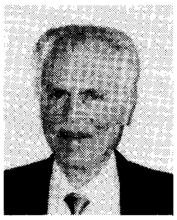
[12] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. New York: McGraw-Hill, 1965.

[13] G. Strang, *Linear Algebra and Applications*. New York: Academic, 1976.

[14] K. S. Miller, *Multidimensional Gaussian Distributions*. New York: Wiley, 1964.

[15] H. Cramer, *Mathematical Methods of Statistics*. Princeton, NJ: Princeton University Press, 1958.

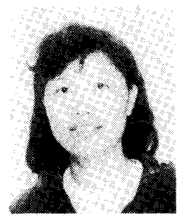
[16] Neyman and Pearson, 1928.



Irving S. Reed (SM'69–F'73) was born in Seattle, WA, on November 12, 1923. He received the B.S. and Ph.D. degrees in mathematics from the California Institute of Technology, Pasadena, in 1944 and 1949, respectively.

From 1951 to 1960 he was associated with Lincoln Laboratory, Massachusetts Institute of Technology, Lexington. From 1960 to 1968 he was a Senior Staff Member with the RAND Corporation, Santa Monica, CA. Since 1963 he has been a Professor of Electrical Engineering and Computer Science at the University of Southern California, Los Angeles. He holds the Charles Lee Powell Professorship in Computer Engineering at USC. He is also a consultant to the RAND Corporation, the MITRE Corporation, and a Director of Adaptive Sensors, Inc. His interests include mathematics, VLSI computer design, coding theory, stochastic process, and information theory.

Dr. Reed is a member of the National Academy of Engineering. In 1989 he received the IEEE Richard W. Hamming medal.



Xiaoli Yu was born in Wuhan, China. She received the M.S. degree and was a Ph.D. candidate in electrical engineering from Tsinghua University, Beijing, China, in 1984 and 1985, respectively, and received the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, in 1990. Her research interests include stochastic process, statistic detection and estimation theory, communication theory, and digital signal processing.