A MAXIMUM LIKELIHOOD DETECTION OF SIGNALS USING FEATURE MAPPING FRAMEWORK

X. Yu†‡, A. M. Chen†

†Communications and Surveillance Division
SAIC
San Diego, CA 92121

I. S. Reed ‡

‡Communications Science Institute
EE Dept.
Los Angeles, CA 90089

ABSTRACT

In [1] a matched-filter based detector was developed for the problem of detecting a 2-D target signal where prior information about the target pattern or template as well as the statistical properties of the clutter is limited. This was accomplished by an *ad hoc* substitution of the maximum likelihood estimate (MLE) of unknown clutter covariance matrix and the MLE's of the complex amplitudes of the significant features components of target into the matched filter test. This paper provides a new approach for the problem based on the generalized likelihood ratio (GLR) principle which maximizes the GLR function over unknown clutter covariance matrix and the unknown significant feature components of target signal to be detected. This new GLR test is compared with the matched-filter based test in [1] for performance. The feature mapping and representation which can be incorporated into the test to characterize the unknown target pattern are various, including the short time Fourier transform, the discrete cosine transform, and the discrete wavelet transform.

1. INTRODUCTION

In most detection applications, clutter covariance matrix and target pattern to be detected are not known apriori with any certainty. If the covariance matrix is unknown for a nonstationary clutter background, then one must account for by using adaptive techniques. One of the adaptive approaches, proposed earlier by Reed, Mallett and Brennan in [2] for not knowing the true clutter covariance matrix, is the *ad hoc* procedure of substituting the maximum likelihood estimate based on the observation data into the test function derived by assuming clutter covariance is known. This method was employed in [1] for the problem to detect a 2-D signal without priori knowing the target pattern as well as the statistical properties of the clutter. The detection test was derived by an *ad hoc* substitution of the maximum likelihood estimate (MLE) of unknown clutter covariance matrix and the MLE's of the complex amplitudes of the significant features components of target into the matched filter test, which assumes a known target pattern with an unknown scale complex amplitude embedded in a complex Gaussian clutter with a known covariance. One may question about the *ad hoc* substitution procedure and be interested in seeing if the GLR test can work better under the same conditions. Therefore the GLR test using feature mapping framework is investigated in this paper to provide another approach for detection of unknown target pattern and adaptation to unknown clutter statistics. This is accomplished by maximizing the GLR function over unknown clutter covariance matrix and the unknown significant feature components of target signal to be detected. The significant features of target for a selected mapping are prior information incorporated into the GLR test to faciliate pattern uncertainty characterization.

This work was funded by ARPA War Breaker Program under Contract N66001-91-D-0046.

2. PROBLEM FORMULATION

A. Data Model: First let

$$\underline{x} = [x_1, x_2, \ldots, x_N]^T \tag{1}$$

denote the N-vector formed by row ordering the pixels of the complex subimage of the observation data. Next let $\underline{s}$ be the complex pattern vector of signal to be detected which is also in a row ordering. Finally, assume that the data vector $\underline{x}$ equals approximately the signal $\underline{s}$, plus a clutter-plus-noise vector $\underline{n}$, i.e.,

$$\underline{x} = \underline{s} + \underline{n}\ , \tag{2}$$

where the vector $\underline{n}$ is assumed to have a complex multivariate normal distribution with zero mean, denoted by $N_{\underline{\mathbf{n}}}(\underline{0}, M)$, under hypotheses H_i, for $i = 0, 1$. Here

0-7803-2431-5/95 $4.00

$M = \mathbf{E}\{\underline{n}\,\underline{n}^h|H_i\}$ denotes the covariance matrix of the clutter background $\underline{n}$ under both H_0 and H_1.

Next, assume that the pixel data surrounding the neighborhood in which the presence of the signal is to be tested is approximately homogeneous (stationary in space). Then let $\underline{y}_l$, for $l = 1, 2, ..., L$, be approximately independent L vectors of data surrounding or in the neighborhood of the signal template $\underline{s}$, which are called "secondary data" and represented by an $N \times L$ matrix $Y = [\underline{y}_1, \underline{y}_2, ..., \underline{y}_L]$. The secondary data matrix Y is subjected to $N_{\mathbf{Y}}(0, I_L \otimes M)$, where I_L is an $L \times L$ identity matrix and $\otimes$ denotes the Kronecker product

B. Feature Mapping and Representation: Let $\underline{s}$ be an N-dimensional signal vector. Then $\underline{s}$ can be represented always by the summation of n linearly independent vectors as follows:

$$\underline{s} = \sum_{n=1}^{N} \underline{\phi}_n b_n = \Phi_K \underline{b}_K + \underline{e} \qquad (3)$$

where

$$\Phi_K = [\underline{\phi}_1 \underline{\phi}_2, ..., \underline{\phi}_K] \quad , \quad \underline{b}_K = [b_1 b_2, ..., b_K]^T$$

and "T" denotes transpose operation. One may call $\underline{\phi}_n$, the n-th *feature* or *feature vector*, and b_n the n-th component of the *feature space*. The matrix Φ_K is assumed to be composed of K most significant features and $\underline{e}$ is the error vector due to the approximation with the K features.

One can use the mean-square magnitude of $\underline{e}(K)$ as the part of a criterion to measure the effectiveness of a subset of K features or basis vectors. If the b_n's are random, it is well-known that the optimal choice for the $\underline{\phi}_n$'s is the eigenvectors of the covariance matrix of $\underline{s}$, i.e. the Karhunen-Loeve expansion. Since the signal covariance matrix is usually unknown and hardly even measured by using observations, the KL expansion is rarely used in practice. Other suboptimal feature mappings may include those from the conventional image transforms to the modern multiresolution decompositions. The traditional image transforms, such as the Fourier Transform and the Discrete Cosine Transform (DCT) provide the frequency domain information of the data. The DCT is excellent on energy compaction for highly correlated data. It is well known that the DCT is a close approximation to a KL transform of the first-order Markov process when the correlation coefficient ρ is close to one. Modern multiresolution decompositions, such as the wavelet and Gabor transforms, are well known for their capability to provide both spatial and frequency localizations.

In the transform feature domain the significant features, i.e., the features which carry most of the signal energy, are extracted with a selected mask. In order to measure the *separability* of the selected target features from the clutter, the locations of the extracted significant target features are compared with the locations of significant clutter features in the transform domain. A large *separability* gives rise to a high signal-to-noise ratio (SNR) and as a consequence has a better detection performance. Here "noise" consists of clutter-plus-receiver noise.

3. MATCHED-FILTER BASED DETECTOR WITH FEATURE MAPPINGS

Consider the classical hypothesis testing given by

$$H_0 : \underline{x} = \underline{n} \qquad H_1 : \underline{x} = \underline{s} + \underline{n}\,. \qquad (4)$$

Let the mean value of the data vector $\underline{x}$ and the covariance matrix of the clutter background $\underline{n}$, under hypotheses H_i, for $i = 0, 1$, be denoted by

$$\mathbf{E}\{\underline{n}|H_0\} = 0, \;\; \mathbf{E}\{\underline{n}|H_1\} = \underline{s}, \;\; M = \mathbf{E}\{\underline{n}\,\underline{n}^h|H_i\}\,. \quad (5)$$

Then the magnitude of the optimal matched filter for a noncoherent detection is given by

$$|y| = |\underline{s}^h M^{-1} \underline{x}| \begin{array}{l} \le \mu \quad \text{then } H_0 \\ > \mu \quad \text{then } H_1 \end{array}\,. \qquad (6)$$

The output of the test criterion in (6) can be made to have a constant false alarm rate (CFAR) by creating the following normalization:

$$u = |y|/\sqrt{\underline{s}^h M^{-1} \underline{s}} = \frac{|\underline{s}^h M^{-1} \underline{x}|}{\sqrt{\underline{s}^h M^{-1} \underline{s}}} \begin{array}{l} \le \rho \quad \text{then } H_0 \\ > \rho \quad \text{then } H_1, \end{array} \qquad (7)$$

where

$$\rho = \mu/\sqrt{\underline{s}^h M^{-1} \underline{s}}\,. \qquad (8)$$

Such a test has a variance of one. The detector in (7) is a maximum invariant test under the group of the change-of-phase and the change-of-scale transformations on $\underline{s}$.

In most applications, the clutter-plus-noise covariance matrix M and the signal vector $\underline{s}$ are not known apriori with any certainty. Thus in such cases, in order to perform the detection test in (7), it is necessary to find statistical estimates of $\hat{M}$ and $\hat{\underline{s}}$ to substitute into the test (7) for M and $\underline{s}$, respectively. A substitution of these estimates into the square of the test in (7) yields the test criterion,

$$r = \frac{|\hat{\underline{s}}^h \hat{M}^{-1} \underline{x}|^2}{\hat{\underline{s}}^h \hat{M}^{-1} \hat{\underline{s}}} \begin{array}{l} \le \quad \rho^2 \text{ then } H_0 \\ > \quad \rho^2 \text{ then } H_1\,. \end{array} \qquad (9)$$

The statistical test in (9) might be called a rational substitute for the optimum invariant test in (7).

It is well known that the maximum likelihood estimates (MLEs) of the covariance matrix M and the unknown components b_k of significant feature vector ϕ_k is given by a matrix form as below, respectively,

$$\hat{M} = \frac{1}{L}\sum_{l=1}^{L} \underline{y}_l \underline{y}_l^h = \frac{1}{L} Y Y^h$$
$$\hat{\underline{b}}_K = (\Phi_K^h \hat{M}^{-1} \Phi_K)^{-1} (\Phi_K^h \hat{M}^{-1} \underline{x}) . \quad (10)$$

A use of Eq. (10) to obtain $\hat{\underline{s}}$ and $\hat{M}$, then sustituting them into (9) yields finally the test,

$$r(\underline{x}, \hat{M}) = \underline{x}^h \hat{M}^{-1} \Phi_K (\Phi_K^h \hat{M}^{-1} \Phi_K)^{-1} \Phi_K^h \hat{M}^{-1} \underline{x} \begin{array}{l} \leq \rho^2 \text{ then } H_0 \ , \\ > \rho^2 \text{ then } H_1 \ . \end{array} \quad (11)$$

4. GENERALIZED LIKELIHOOD RATIO BASED DETECTOR WITH FEATURE MAPPINGS

Under the data model assumption and target feature representation described in Sec. 2, the two hypotheses which the detector must distinguish can be alternatively formulated by

$$H_0 : \begin{cases} \underline{x} = \underline{n} \\ \underline{y}_1 = \underline{n}_1 \\ \vdots \\ \underline{y}_L = \underline{n}_L \end{cases} \qquad H_1 : \begin{cases} \underline{x} = \underline{s} + \underline{n} = \Phi_K \underline{b}_K + \underline{n} \\ \underline{y}_1 = \underline{n}_1 \\ \vdots \\ \underline{y}_L = \underline{n}_L \end{cases}$$

where the joint probability density functions of $\underline{x}, Y$ under hypotheses H_i, for $i = 0, 1$, are given as follows:

$$\begin{aligned} p(\underline{x}, Y|H_0) &= p(\underline{x}|H_0) p(Y|H_0) = \frac{1}{(\pi)^{N(L+1)} |M|^{(L+1)}} \\ & \quad \exp\{-(\underline{x}^h M^{-1} \underline{x} + \sum_{l=1}^{L} \underline{y}_l^h M^{-1} \underline{y}_l)\} \\ p(\underline{x}, Y|H_1) &= p(\underline{x}|H_1) p(Y|H_1) = \frac{1}{(\pi)^{N(L+1)} |M|^{(L+1)}} \\ & \quad \exp\{-[(\underline{x} - \Phi_K \underline{b}_K)^h M^{-1} (\underline{x} - \Phi_K \underline{b}_K) \\ & \quad + \sum_{l=1}^{L} \underline{y}_l^h M^{-1} \underline{y}_l)\} . \end{aligned} \quad (12)$$

Then the generalized maximum likelihood ratio test is formulated by:

$$\Lambda(\underline{x}, Y) = \frac{\max_{\underline{b}_K, M} p(\underline{x}, Y|H_1, \underline{b}_K, M)}{\max_M p(\underline{x}, Y|H_0, M)} \begin{array}{l} \geq \Lambda_0 \text{ then } H_1 \\ < \Lambda_0 \text{ then } H_0 \end{array} . \quad (13)$$

Maximizing the numerator and denominator in (13) with respect to the unknown parameters yields the testing function $\Lambda(\underline{x}, Y)$ as follows:

$$\lambda(\underline{x}, Y) = \frac{\underline{x}^h \hat{M}^{-1} \Phi_K (\Phi_K^h \hat{M}^{-1} \Phi_K)^{-1} \Phi_K^h \hat{M}^{-1} \underline{x}}{1 + \frac{1}{L} \underline{x}^h \hat{M}^{-1} \underline{x}} \begin{array}{l} \leq \lambda_0 \text{ then } H_0 \ , \\ > \lambda_0 \text{ then } H_1 \ . \end{array} \quad (14)$$

where $\lambda_0 = L(1 - \frac{1}{\Lambda_0})$ and $\hat{M}$ is defined in Eq. (10) The relationship of this GLR test to the matched filter based test is revealed by the following equation,

$$\lambda(\underline{x}, \hat{M}) = \frac{r(\underline{x}, \hat{M})}{1 + \frac{1}{L} \underline{x}^h \hat{M}^{-1} \underline{x}} \quad (15)$$

This explicit expression indicates that when L becomes larger, two tests tend to be the same.

5. EXPERIMENTAL RESULTS USING ACTUAL SAR DATA

The new developed GLR detector is compared with the match filter based test by using the actual synthetic aperture radar (SAR) data collected by by MIT Lincoln Laboratory on the SRI Ultra Wideband (UWB) SAR sensor. The observed image data, called Illustrative-Example image, is shown in Fig. 1, which consists of twelve military vehicles (three are camouflaged) and indicated by T_1 through T_{12}.

The detection results obtained by applying the new GLR test and the match filter based test are shown in Fig. 2 and 3 respectively, where the discrete cosine transform is used for target feature mapping and representation. The false detections are indicated by F_1 through F_3. A comparison of two resulting detection images reveals that for the same number of correct detections, the number of false alarms caused by the match filter based test is a little bit larger than by that of the GLR test. However, the computation complexity of the match filter based test is lower than the GLR test. An analytic performance comparison is under investigation.

6. REFERENCES

[1] X. Yu and I. S. Reed "Adaptive Detection of Signals with Linear Feature Mappings and Representations", submitted to IEEE Trans. on Signal Processing, 1993.

[2] I. S. Reed, Mallett, S. D. and Brennan, L E. "Rapid convergence rate in adaptive arrays", IEEE Trans. on Aerospace and Electronic Systems, AES-10, 6, Nov. 1974.

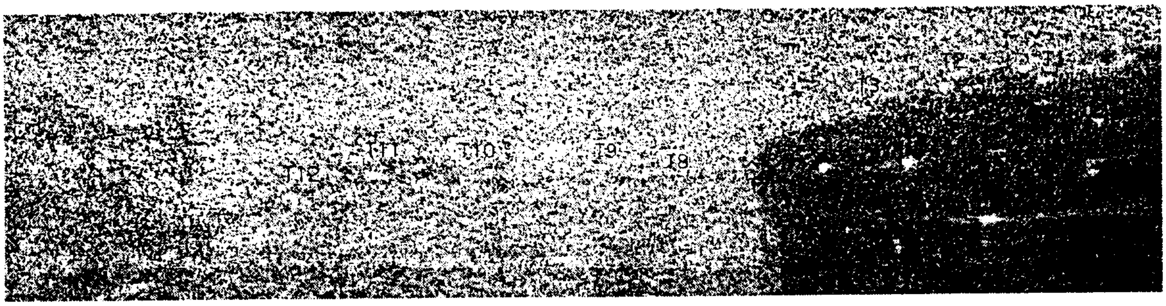

Figure 1: Illustrative-Example SAR image of the SRI FOPEN II ultra-wideband SAR system with HH-polarization, and a 1 meter resolution.

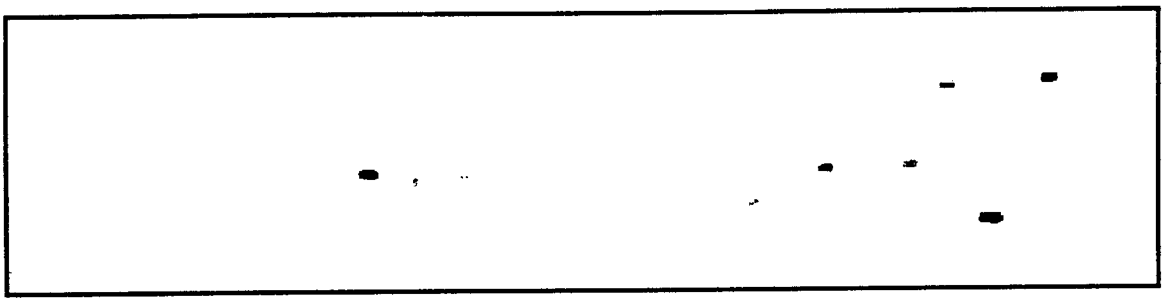

Figure 2: Automatic target detection results by applying the GLR test to the Illustrative-Example Site of Fig. 1. The results are obtained by using DCT linear feature map with the number of significant features being $K = 16$ out of 128.

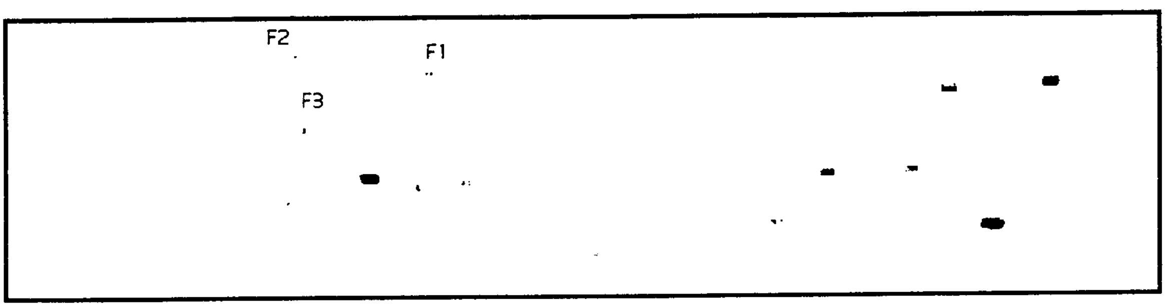

Figure 3: Automatic target detection results by applying the match filter based test to the Illustrative-Example Site of Fig. 1. The results are obtained by using DCT linear feature map with the number of significant features being $K = 16$ out of 128.