Proceedings of the SPIE Conference on Automatic Target Recognition VI **4379,** April, 2001

SIGNATURE ADAPTIVE MINE DETECTION AT A CONSTANT FALSE ALARM RATE

Frank Crosby*, Stephen Riley+
*Veridian Systems, Panama City, FL 32407
+Mechanical Simulation Corporation, Ann Arbor, MI 48103

ABSTRACT

A constant false alarm rate (CFAR) algorithm has been developed for use in multi-band mine detection. While it is often difficult to predict the spectral signatures of targets, the shape of the target may be known. This test exploits geometric target features and spectral differences between the target and the surrounding area. The algorithm is derived from a general statistical model of the data, which allows it to adapt to changing backgrounds and variable target signatures.

Keywords: ATR, CFAR**,** multispectral detection**.**

1. INTRODUCTION

The pervasiveness of land mines and unexploded ordnance has kept their detection at the forefront of remote sensing application research. Recently, multispectral sensors have been developed to aid in their detection[1]. Low false alarm rate algorithms can be used with multispectral images to improve clearance and avoidance of mines and mine-like objects. The constant false alarm rate (CFAR) algorithm described in this paper allows for a pre-computed threshold based on desired performance and not on changing scene characteristics.

Multispectral processing exploits spectral energy differences between targets and their surroundings. It has been shown that multi-band processing always enhances the signal to noise ratio available for detection when compared to single band processing. However, the exploitation of spectral differences is complicated by highly variable target and background signatures. The number and wavelengths of the bands which contribute most to detection are also highly variable. Adaptive algorithms owe much of their success to the few assumptions placed on the signatures. A limited number of assumptions allows them to be successful in many different situations.

In addition to their spectral characteristics, most objects of interest are man-made and exhibit highly geometrical structures. Unlike spectral signatures, the shape of the target can often be specified with near certainty.

The algorithm presented in this paper uses both spectral differences and shape information. It is derived from a small number of statistical assumptions on the data. Its detection threshold depends only on the number of spectral bands and the number of pixels in samples. Since the threshold only depends on these two numbers, it is said to detect targets at a constant false alarm rate. Furthermore, it adapts to a changing environment and target signatures by estimating conditions based on the observed data.

The following section describes the statistical model and develops the algorithm from the assumptions. The third section shows that the probability of a false alarm (PFA) for the test can be found using a standard distribution with the appropriate parameters. That is followed by a demonstration of the algorithm's performance on actual data. The results show that if the data follows the initial statistical model, then the resulting test is very effective.

2. STATISTICAL MODEL

Previous studies have found that multispectral optical images can be locally modeled as independent Gaussian processes[2]. Performing local analysis reduces the possibility of encountering samples with several spectral distributions. The goal is to analyze an area that contains either a single spectral category or a single category and a target. The local area surrounding the possible target area is modeled with a normal distribution. The target is also assumed to have a normal distribution and share a common covariance matrix with the surrounding area.

An example of a true distribution is shown in figure 1, which is an image of one band from a six-band multispectral image obtained from the Coastal Systems Station, Dahlgren Division Naval Surface Warfare Center, Panama City Florida. The image was collected with an intensified, multispectral sensor that uses a spinning filter wheel in front of the intensifier. A region of interest has been selected near a target. Figure 2 shows a histogram of that area, which demonstrates an approximately Gaussian distribution.

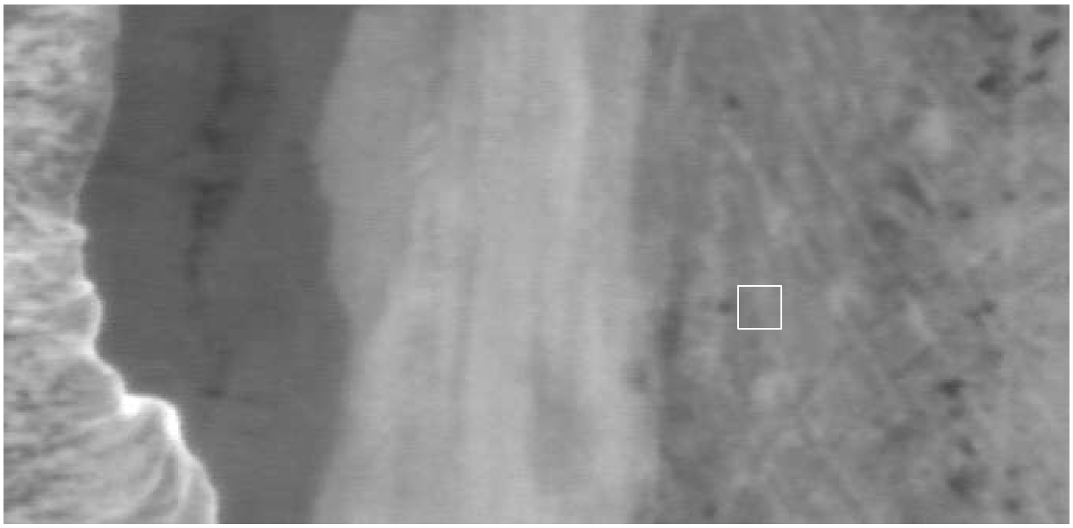

Figure 1. Single band from multispectral image with highlighted region of interest

The initial assumptions of some models require local mean removal to achieve a zero mean distribution[3]. Then if the target is present, the mean of the area will be the target mean. In mine detection applications, the mine resolution can have an effect on the local mean calculation. If the mean of the target is added to the area mean prior to the subtraction, it is possible that the area will have a zero mean regardless of whether a target is present or not.

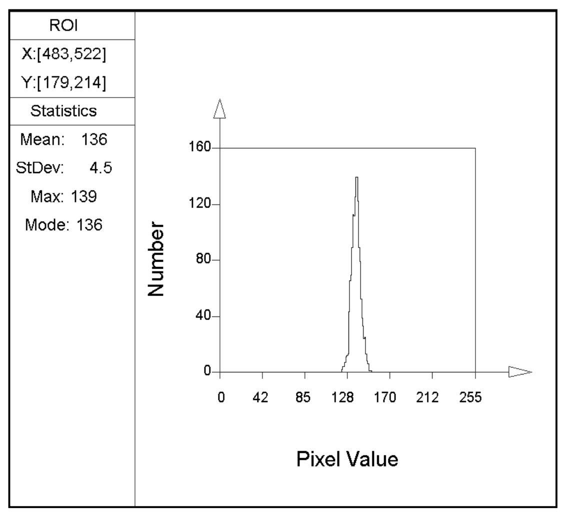

Figure 2. Histogram of region in fig. 1

We wish to determine if an area contains, as a subset, an object of interest. The null hypothesis is that there is no object present. If there is no object present, then the mean of the surrounding area will be similar to the mean of the possible target area. The alternative hypothesis is that there is an object present. If there is an object, then the two means will be different.

In statistical terms, the distributions are assumed to be normal with means $\mu_{\mathbf{B}}$ and $\mu_{\mathbf{T}}$ and common covariance Σ, denoted $N(\mu_{\mathbf{B}}, \Sigma)$ and $N(\mu_{\mathbf{T}}, \Sigma)$. The hypotheses are:

$$H_0 : \mu_{\mathbf{B}} = \mu_{\mathbf{T}}$$
$$H_1 : \mu_{\mathbf{B}} \neq \mu_{\mathbf{T}}$$

In p-band multispectral images, each pixel is a p-dimensional sample, $\mathbf{x}_i = [x_1 \cdots x_p]^{\mathrm{T}}$. A local sample of the image of size N is given by $\mathbf{X} = [\mathbf{x}_1 \cdots \mathbf{x}_N]$. This sample is divided into two disjoint sets, background pixels, $\mathbf{B} = [\mathbf{b}_0, \ldots, \mathbf{b}_{N_{\mathbf{B}}}]$, and possible target pixels, $\mathbf{T} = [\mathbf{t}_0, \ldots, \mathbf{t}_{N_{\mathbf{T}}}]$. Since this local area contains only background and possible target pixels,

$N_{\mathbf{B}} + N_{\mathbf{T}} = N$. Let $\mathbf{M_B}$ be a column vector of length N satisfying $\mathbf{M_B M_B^T} = 1$, whose nonzero entries denote the pixels in the background. The mean of the background pixels is given by $\mathbf{XM_B}$. In the derivation of the hypothesis test, it will be useful to refer to the individual pixels in the background. These shall be denoted by $\mathbf{b}_i$, $i = 1, \ldots, N$. Similarly, let $\mathbf{M_T}$ be a vector of length N satisfying $\mathbf{M_T M_T^T} = 1$, whose nonzero entries denote the pixels in the possible target area. The mean of the possible target area is given by $\mathbf{XM_T}$. The individual pixels will be denoted by $\mathbf{t}_i$, $i = 1, \ldots, N$.

In order to reliably test the hypothesis we require some knowledge of the distributions $N(\mu_{\mathbf{B}}, \Sigma)$ and $N(\mu_{\mathbf{T}}, \Sigma)$. In most practical applications the exact mean and covariance of the background and target are unknown. The assumed distributions and the samples may be used to elicit estimators for these parameters. Maximum likelihood estimators provide good estimates. Although there are several different methods for estimating unknown statisical parameters under fairly general conditions maximum likelihood estimators satisfy some desirable properties. The maximum likelihood estimator has a normal distribution with a mean equal to the true parameter with a minimum variance[4].

The maximum likelihood estimators are found by maximizing the likelihood function, which is the joint probability function of the individual pixels. The maximum likelihood method chooses estimates of the unknown parameters which maximize the likelihood function. This can often be accomplished by finding critical points through differentiation. When the likelihood function involves an exponential term, it can be more convenient to find the critical points of the natural logarithm of the likelihood function. The log function is monotonically increasing, and so a point that maximizes the log of the likelihood function also maximizes the likelihood function.

The estimation of these parameters is used in constructing the test of the hypothesis. If the actual parameters are known, then they are used instead.

The likelihood function for a background and target sample is given by

$$L(\mu_{\mathbf{B}}, \mu_{\mathbf{T}}, \Sigma\ ) = \left(\frac{1}{(2\pi)^p |\Sigma\ |^p}\right)^{N_{\mathbf{B}}/2} \exp\left(\frac{-1}{2}\sum_{i=0}^{N}(\mathbf{b}_i - \mu_{\mathbf{B}})^{\mathrm{T}} \Sigma^{-1}(\mathbf{b}_i - \mu_{\mathbf{B}})\right)$$

$$\cdot\left(\frac{1}{(2\pi)^p |\Sigma\ |^p}\right)^{N_{\mathbf{T}}/2} \exp\left(\frac{-1}{2}\sum_{j=0}^{N_{\mathbf{T}}}(\mathbf{t}_j - \mu_{\mathbf{T}})^{\mathrm{T}} \Sigma^{-1}(\mathbf{t}_j - \mu_{\mathbf{T}})\right)$$

$$= \left((2\pi)^p\right)^{-(N)/2} \left(|\Sigma\ |^{-1}\right)^{N/2} \exp\left(\frac{-1}{2}\left(\sum_{i=1}^{N_{\mathbf{B}}}(\mathbf{b}_i - \mu_{\mathbf{B}})^{\mathrm{T}} \Sigma^{-1}(\mathbf{b}_i - \mu_{\mathbf{B}}) + \sum_{i=1}^{N_{\mathbf{T}}}(\mathbf{t}_i - \mu_{\mathbf{T}})^{\mathrm{T}} \Sigma^{-1}(\mathbf{t}_i - \mu_{\mathbf{T}})\right)\right),$$

which is also the joint probability distribution function. We shall differentiate

$$\ln(L) = \frac{-(N_{\mathbf{B}} + N_{\mathbf{T}})}{2}\left[\ln[(2\pi)^p] + \ln|\Sigma^{-1}|\right]$$

$$-\left(\frac{1}{2}\right)\left[\sum_{i=1}^{N_{\mathbf{B}}}(\mathbf{b}_i - \mu_{\mathbf{B}})^{\mathrm{T}} \Sigma^{-1}(\mathbf{b}_i - \mu_{\mathbf{B}}) + \sum_{j=1}^{N_{\mathbf{T}}}(\mathbf{t}_j - \mu_{\mathbf{T}})^{\mathrm{T}} \Sigma^{-1}(\mathbf{t}_j - \mu_{\mathbf{T}})\right].$$

Let $s \in \sum^{-1}$, then

$$\ln(L) = \frac{-N}{2} p\ln(2\pi) + \frac{N}{2}\ln|\Sigma^{-1}|$$

$$-\left(\frac{1}{2}\right)\left[\sum_{i=0}^{N_{\mathbf{B}}}\left[\sum_{l,k=1}^{p} s_{lk}(\mathbf{b}_i - \mu_{\mathbf{B}})_l(\mathbf{b}_i - \mu_{\mathbf{B}})_k\right] + \sum_{i=1}^{N_{\mathbf{T}}}\left[\sum_{l,k=0}^{p} s_{lk}(\mathbf{t}_i - \mu_{\mathbf{T}})_l(\mathbf{t}_i - \mu_{\mathbf{T}})_k\right]\right].$$

The differentiation uses the following identities,

$\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} = \sum_{i,j=1}^{p} a_{ij} x_i x_j$, $\frac{\partial}{\partial \mathbf{x}}\left(\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}\right) = 2\mathbf{A}\mathbf{x}$, and $\frac{\partial}{\partial s_{lk}}\left|\Sigma\right| = \Sigma_{lk}$, where Σ_{lk} is the cofactor of s_{lk} .

The null hypothesis defines a subset of the parameter space. Let ω be the set of points in the parameter space which correspond to H_0 and Ω be the set of all points in the parameter space. For our example, $\omega = \{(\mu_{\mathbf{B}}, \mu_{\mathbf{T}}) : \mu_{\mathbf{B}} = \mu_{\mathbf{T}}, -\infty < \mu_{\mathbf{B}}, \mu_{\mathbf{T}} < \infty\}$ and $\Omega = \{(\mu_{\mathbf{B}}, \mu_{\mathbf{T}}) : -\infty < \mu_{\mathbf{B}}, \mu_{\mathbf{T}} < \infty\}$.

If $\mu_{\mathbf{B}}, \mu_{\mathbf{T}}$ and Σ are in ω, then $\mu_{\mathbf{B}} - \mu_{\mathbf{T}} = 0$. We denote the common mean $\mu = \mu_{\mathbf{B}} = \mu_{\mathbf{T}}$ and find

$\frac{\partial}{\partial \mu}\ln(L) = \frac{-1}{2}\left[\sum_{i=1}^{N} 2\,\Sigma^{-1}(\mathbf{b}_i - \mu) + \sum_{j=1}^{N_{\mathbf{T}}} 2\,\Sigma^{-1}(\mathbf{t}_i - \mu)\right]$. The maximum likelihood estimate of the mean is given by

$$\hat{\mu} = \frac{\sum_{i=1}^{N_{\mathbf{B}}}\mathbf{b}_i + \sum_{i=1}^{N_{\mathbf{T}}}\mathbf{t}_i}{N_{\mathbf{B}} + N_{\mathbf{T}}} = \left(\frac{1}{N_{\mathbf{B}} + N_{\mathbf{B}}}\right)\left(\frac{\mathbf{XM}_{\mathbf{B}}}{N_{\mathbf{B}}} + \frac{\mathbf{XM}_{\mathbf{T}}}{N_{\mathbf{T}}}\right).$$

Differentiating with respect to an element, s, in Σ^{-1} gives

$\frac{\partial}{\partial s_{lk}}\ln(L) = \frac{-N}{2}\left(\frac{\Sigma^{-1}}{\left|\Sigma^{-1}\right|}\right) - \left(\frac{1}{2}\right)\left[\sum_{i=1}^{N_{\mathbf{B}}}(\mathbf{b}_i - \mu_{\mathbf{B}})(\mathbf{b}_i - \mu_{\mathbf{B}})^{\mathrm{T}} + \sum_{i=1}^{N_{\mathbf{T}}}(\mathbf{t}_i - \mu_{\mathbf{T}})(\mathbf{t}_i - \mu_{\mathbf{T}})^{\mathrm{T}}\right]$. To simplify $\frac{\partial}{\partial s_{lk}}\ln(L)$, we use the identity for the (l, k)-th entry of Σ, $\sigma_{lk} = \frac{\Sigma_{lk}^{-1}}{\left|\Sigma^{-1}\right|}$. Thus,

$\frac{\partial}{\partial s_{lk}}\ln(L) = \frac{-N}{2}\sigma_{lk} - \frac{1}{2}\left[\sum_{i=1}^{N_{\mathbf{B}}}(\mathbf{b}_i - \mu_{\mathbf{B}})(\mathbf{b}_i - \mu_{\mathbf{B}})^{\mathrm{T}} + \sum_{i=1}^{N_{\mathbf{T}}}(\mathbf{t}_i - \mu_{\mathbf{T}})(\mathbf{t}_i - \mu_{\mathbf{T}})^{\mathrm{T}}\right]$ and

$\hat{\sigma}_{lk} = \frac{1}{N}\left[\sum_{i=1}^{N_{\mathbf{B}}}(\mathbf{b}_i - \mu_{\mathbf{B}})(\mathbf{b}_i - \mu_{\mathbf{B}})^{\mathrm{T}} + \sum_{i=1}^{N_{\mathbf{T}}}(\mathbf{t}_i - \mu_{\mathbf{T}})(\mathbf{t}_i - \mu_{\mathbf{T}})^{\mathrm{T}}\right]_{lk}$. The maximum likelihood estimate of the covariance is given by

$$\begin{aligned}\hat{\Sigma} &= \frac{1}{N}\left[\sum_{i=1}^{N_{\mathbf{B}}}(\mathbf{b}_i - \mu_{\mathbf{B}})(\mathbf{b}_i - \mu_{\mathbf{B}})^{\mathrm{T}} + \sum_{i=1}^{N_{\mathbf{T}}}(\mathbf{t}_i - \mu_{\mathbf{T}})(\mathbf{t}_i - \mu_{\mathbf{T}})^{\mathrm{T}}\right] \\ &= \frac{1}{N}\left[(\mathbf{B} - \mathbf{XM}_{\mathbf{B}})(\mathbf{B} - \mathbf{XM}_{\mathbf{B}})^{\mathrm{T}} + (\mathbf{T} - \mathbf{XM}_{\mathbf{T}})(\mathbf{T} - \mathbf{XM}_{\mathbf{T}})^{\mathrm{T}}\right].\end{aligned}$$

Finding the values of the variables which maximize L subject to their membership in Ω, we find that the maximum likelihood estimates, $\hat{\mu}_{\mathbf{B}}, \hat{\mu}_{\mathbf{T}}$ and $\hat{\Sigma}$ are

$\hat{\mu}_{\mathbf{B}} = \frac{1}{N_{\mathbf{B}}}\sum_{i=1}^{N}\mathbf{b}_i = \mathbf{XM}_{\mathbf{B}}$, $\hat{\mu}_{\mathbf{T}} = \frac{1}{N_{\mathbf{T}}}\sum_{i=1}^{N_{\mathbf{T}}}\mathbf{t}_i = \mathbf{XM}_{\mathbf{T}}$, and $\hat{\Sigma}$ remains unchanged.

Establishing a test with a constant false alarm rate is based on the probability of rejecting H_0 when H_0 is true. If α is the desired false alarm rate, then the level of test may be written

$P(\,(\mathbf{B},\mathbf{T})\in\Omega\setminus\omega\,|\,H_0 \text{ true})\le\alpha$. The set of points in the parameter space which corresponds to a rejection of the null hypothesis is called the critical region. The critical region depends on both the joint probability density function and the desired false alarm rate.

By utilizing the likelihood function we can judge if a particular sample of the background and target does not support H_0. Let the sample be given and let $L(\hat{\omega})$ be the likelihood function maximized with respect to parameters in ω, and let $L(\hat{\Omega})$ be the likelihood function maximized with respect to parameters in Ω. One would expect that if the sample does not support H_0, then $L(\hat{\omega})$ would be much smaller than $L(\hat{\Omega})$. This was formally proven by Neyman-Pearson in their generalized maximum likelihood ratio criterion. The criterion states that the critical region for a test of H_0 against H_1 is defined by the set of points in the sample space for which $\frac{L(\hat{\omega})}{L(\hat{\Omega})}\le k$, where k is selected so that $P(\lambda<k\,|\,H_0 \text{ true})\le\alpha$.

Placing the maximum likelihood estimates into the likelihood function gives

$$L(\hat{\Omega})=\left[\frac{1}{(2\pi)^p\left|\frac{\sum_{i=1}^{N_\mathbf{B}}(\mathbf{b}_i-\hat{\mu}_\mathbf{B})(\mathbf{b}_i-\hat{\mu}_\mathbf{B})^\mathrm{T}+\sum_{i=1}^{N_\mathbf{T}}(\mathbf{t}_i-\hat{\mu}_\mathbf{T})(\mathbf{t}_i-\hat{\mu}_\mathbf{T})^\mathrm{T}}{N_\mathbf{B}+N_\mathbf{T}}\right|}\right]^{\frac{N}{2}}$$
$$\times\exp\left[\sum_{i=1}^{N_\mathbf{B}}(\mathbf{b}_i-\hat{\mu}_\mathbf{B})^\mathrm{T}\,\hat{\Sigma}^{-1}(\mathbf{b}-\hat{\mu}_\mathbf{B})+\sum_{i=1}^{N_\mathbf{T}}(\mathbf{t}_i-\hat{\mu}_\mathbf{T})^\mathrm{T}\,\hat{\Sigma}^{-1}(\mathbf{t}_i-\hat{\mu}_\mathbf{T})\right]$$

Since $\sum_{i=1}^{N_\mathbf{B}}(\mathbf{b}-\hat{\mu}_\mathbf{B})^\mathrm{T}\,\hat{\Sigma}^{-1}(\mathbf{b}-\hat{\mu}_\mathbf{B})$ and $\sum_{i=1}^{N_\mathbf{T}}(\mathbf{t}_i-\hat{\mu}_\mathbf{T})^\mathrm{T}\,\hat{\Sigma}^{-1}(\mathbf{t}_i-\hat{\mu}_\mathbf{T})$ are scalers, the properties of the trace of a matrix allow for the following simplification.

$$\begin{aligned}\sum_{i=1}^{N_\mathbf{B}}(\mathbf{b}-\hat{\mu}_\mathbf{B})^\mathrm{T}\,\hat{\Sigma}^{-1}(\mathbf{b}-\hat{\mu}_\mathbf{B})+\sum_{i=1}^{N_\mathbf{T}}(\mathbf{t}_i-\hat{\mu}_\mathbf{T})^\mathrm{T}\,\hat{\Sigma}^{-1}(\mathbf{t}_i-\hat{\mu}_\mathbf{T}) &= \mathrm{tr}\left\{\sum_{i=1}^{N_\mathbf{B}}(\mathbf{b}_i-\hat{\mu}_\mathbf{B})^\mathrm{T}\,\hat{\Sigma}^{-1}(\mathbf{b}_i-\hat{\mu}_\mathbf{B})+\sum_{i=1}^{N_\mathbf{T}}(\mathbf{t}_i-\hat{\mu}_\mathbf{T})^\mathrm{T}\,\hat{\Sigma}^{-1}(\mathbf{t}_i-\hat{\mu}_\mathbf{T})\right\}\\ &= \mathrm{tr}\left\{\sum_{i=1}^{N_\mathbf{B}}\hat{\Sigma}^{-1}(\mathbf{b}_i-\hat{\mu}_\mathbf{B})(\mathbf{b}_i-\hat{\mu}_\mathbf{B})^\mathrm{T}+\sum_{i=1}^{N_\mathbf{T}}\hat{\Sigma}^{-1}(\mathbf{t}_i-\hat{\mu}_\mathbf{T})(\mathbf{t}_i-\hat{\mu}_\mathbf{T})^\mathrm{T}\right\}\\ &= \mathrm{tr}\left\{\sum_{i=1}^{N_\mathbf{B}}\mathbf{I}_p+\sum_{i=1}^{N_\mathbf{T}}\mathbf{I}_p\right\}\\ &= N_\mathbf{B}\,p+N_\mathbf{T}\,p\end{aligned}$$

Therefore,

$$\exp\left[\sum_{i=1}^{N_\mathbf{B}}(\mathbf{b}_i-\hat{\mu}_\mathbf{B})^\mathrm{T}\,\hat{\Sigma}^{-1}(\mathbf{b}_i-\hat{\mu}_\mathbf{B})+\sum_{i=1}^{N_\mathbf{T}}(\mathbf{t}_i-\hat{\mu}_\mathbf{T})^\mathrm{T}\,\hat{\Sigma}^{-1}(\mathbf{t}_i-\hat{\mu}_\mathbf{T})\right]=\exp\left[\frac{-(N_\mathbf{B}+N_\mathbf{T})p}{2}\right]$$ and

$$L(\hat{\Omega}) = \left[\frac{(N_{\mathbf{B}} + N_{\mathbf{T}})pe^{-1}}{(2\pi)^p \left| \sum_{i=1}^{N_{\mathbf{B}}} (\mathbf{b}_i - \hat{\mu}_{\mathbf{B}})(\mathbf{b}_i - \hat{\mu}_{\mathbf{B}})^{\mathrm{T}} + \sum_{j=1}^{N_{\mathbf{T}}} (\mathbf{t}_i - \hat{\mu}_{\mathbf{T}})(\mathbf{t}_i - \hat{\mu}_{\mathbf{T}})^{\mathrm{T}} \right|} \right]^{\frac{N}{2}}$$

Similarly,

$$L(\hat{\omega}) = \left[\frac{(N_{\mathbf{B}} + N_{\mathbf{T}})pe^{-1}}{(2\pi)^p \left| \sum_{i=1}^{N_B} (\mathbf{b}_i - \hat{\mu})(\mathbf{b}_i - \hat{\mu})^{\mathrm{T}} + \sum_{j=1}^{N_{\mathbf{T}}} (\mathbf{t}_i - \hat{\mu})(\mathbf{t}_i - \hat{\mu})^{\mathrm{T}} \right|} \right]^{\frac{N}{2}}$$

The likelihood ratio is

$$d(\mathbf{X}) = \frac{L(\hat{\omega})}{L(\hat{\Omega})} = \left[\frac{(N_{\mathbf{B}} + N_{\mathbf{T}})pe^{-1}}{(2\pi)^p \left| \sum_{i=1}^{N_{\mathbf{B}}} (\mathbf{b}_i - \hat{\mu}\ )^{\mathrm{T}}(\mathbf{b}_i - \hat{\mu}\ )^{\mathrm{T}} + \sum_{j=1}^{N_{\mathbf{T}}} (\mathbf{t}_i - \hat{\mu}\ )(\mathbf{t}_i - \hat{\mu}\ )^{\mathrm{T}} \right|} \right]^{\frac{N}{2}}$$

$$\times \left[\frac{(2\pi)^p \left| \sum_{i=1}^{N_{\mathbf{B}}} (\mathbf{b}_i - \hat{\mu}_{\mathbf{X}})(\mathbf{b}_i - \hat{\mu}_{\mathbf{X}})^{\mathrm{T}} + \sum_{j=1}^{N_{\mathbf{T}}} (\mathbf{t}_j - \hat{\mu}_{\mathbf{Y}})(\mathbf{t}_j - \hat{\mu}_{\mathbf{Y}})^{\mathrm{T}} \right|}{(N_{\mathbf{B}} + N_{\mathbf{T}})pe^{-1}} \right]^{\frac{N}{2}}$$

$$= \left[\frac{\left| \sum_{i=1}^{N_{\mathbf{B}}} (\mathbf{b}_i - \mu_{\mathbf{B}})(\mathbf{b}_i - \mu_{\mathbf{B}})^{\mathrm{T}} + \sum_{i=1}^{N_{\mathbf{T}}} (\mathbf{t}_i - \hat{\mu}_{\mathbf{T}})(\mathbf{t}_i - \hat{\mu}_{\mathbf{T}})^{\mathrm{T}} \right|}{\left| \sum_{i=1}^{N_{\mathbf{B}}} (\mathbf{b}_i - \hat{\mu}\ )(\mathbf{b}_i - \hat{\mu}\ )^{\mathrm{T}} + \sum_{i=1}^{N_{\mathbf{T}}} (\mathbf{t}_i - \hat{\mu}\ )(\mathbf{t}_i - \hat{\mu}\ )^{\mathrm{T}} \right|} \right]^{\frac{N}{2}}$$

$$= \left[\frac{\left| \sum_{i=1}^{N_{\mathbf{B}}} (\mathbf{b}_i - \hat{\mu}_{\mathbf{B}})(\mathbf{b}_i - \hat{\mu}_{\mathbf{B}})^{\mathrm{T}} + \sum_{i=1}^{N_{\mathbf{T}}} (\mathbf{t}_i - \hat{\mu}_{\mathbf{T}})(\mathbf{t}_i - \hat{\mu}_{\mathbf{T}})^{\mathrm{T}} \right|}{\left| \sum_{i=1}^{N_{\mathbf{B}}} (\mathbf{b}_i - \hat{\mu}_{\mathbf{B}})(\mathbf{b}_i - \hat{\mu}_{\mathbf{B}})^{\mathrm{T}} + \sum_{j=1}^{N_{\mathbf{T}}} (\mathbf{t}_j - \hat{\mu}_{\mathbf{T}})(\mathbf{t}_j - \hat{\mu}_{\mathbf{T}})^{\mathrm{T}} + N_{\mathbf{B}}(\hat{\mu}_{\mathbf{B}} - \hat{\mu})^{\mathrm{T}}(\hat{\mu}_{\mathbf{B}} - \hat{\mu}) + N_{\mathbf{T}}(\hat{\mu}_{\mathbf{T}} - \hat{\mu})^{\mathrm{T}}(\hat{\mu}_{\mathbf{T}} - \hat{\mu}) \right|} \right]^{\frac{N}{2}}$$

Simiplifying the last term in the denominator gives

$$N_{\mathbf{B}}(\hat{\mu}_{\mathbf{X}} - \hat{\mu})^{\mathrm{T}}(\hat{\mu}_{\mathbf{X}} - \hat{\mu}) + N_{\mathbf{T}}(\hat{\mu}_{\mathbf{Y}} - \hat{\mu})^{\mathrm{T}}(\hat{\mu}_{\mathbf{Y}} - \hat{\mu})$$

$$= N_{\mathbf{B}} \left(\frac{\sum \mathbf{b}_i}{N_{\mathbf{B}}} - \frac{\sum \mathbf{b}_i + \sum \mathbf{t}_i}{N} \right) \left(\frac{\sum \mathbf{b}_i}{N_{\mathbf{B}}} - \frac{\sum \mathbf{b}_i + \sum \mathbf{t}_i}{N} \right) + N_{\mathbf{T}} \left(\frac{\sum \mathbf{t}_i}{N_{\mathbf{T}}} - \frac{\sum \mathbf{b}_i + \sum \mathbf{t}_i}{N} \right) \left(\frac{\sum \mathbf{t}_i}{N_{\mathbf{T}}} - \frac{\sum \mathbf{b}_i + \sum \mathbf{t}_i}{N} \right)$$

$= \frac{N_{\mathbf{B}} N_{\mathbf{T}}}{N} (\hat{\mu}_{\mathbf{B}} - \hat{\mu}_{\mathbf{T}})(\hat{\mu}_{\mathbf{B}} - \hat{\mu}_{\mathbf{T}})$. Placing this term in the likelihood ratio gives

$$\frac{L(\hat{\Omega})}{L(\hat{\omega})} = \left[\frac{\left| \sum_{i=1}^{N_{\mathbf{B}}} (\mathbf{b}_i - \hat{\mu}_{\mathbf{B}})(\mathbf{b}_i - \hat{\mu}_{\mathbf{B}})^{\mathrm{T}} + \sum_{i=1}^{N_{\mathbf{T}}} (\mathbf{t}_i - \hat{\mu}_{\mathbf{T}})(\mathbf{t}_i - \hat{\mu}_{\mathbf{T}})^{\mathrm{T}} \right|}{\left| \sum_{i=1}^{N_{\mathbf{B}}} (\mathbf{b}_i - \hat{\mu}_{\mathbf{B}})(\mathbf{b}_i - \hat{\mu}_{\mathbf{B}})^{\mathrm{T}} + \sum_{i=1}^{N_{\mathbf{T}}} (\mathbf{t}_i - \hat{\mu}_{\mathbf{T}})(\mathbf{t}_i - \hat{\mu}_{\mathbf{T}})^{\mathrm{T}} + \frac{N_{\mathbf{B}} N_{\mathbf{T}}}{N} (\hat{\mu}_{\mathbf{B}} - \hat{\mu}_{\mathbf{T}})^{\mathrm{T}} (\hat{\mu}_{\mathbf{B}} - \hat{\mu}_{\mathbf{T}}) \right|} \right]^{\frac{N}{2}}$$

. Using the determinate identity, $\begin{vmatrix} \mathbf{C} & \mathbf{D} \\ \mathbf{E} & \mathbf{F} \end{vmatrix} = |\mathbf{C}| \bullet |\mathbf{F} - \mathbf{E}\mathbf{C}^{-1}\mathbf{D}|$, the denominator may be written as

$$\left| \sum_{i=1}^{N_{\mathbf{B}}} (\mathbf{b}_i - \hat{\mu}_{\mathbf{B}})(\mathbf{b}_i - \hat{\mu}_{\mathbf{B}})^{\mathrm{T}} + \sum_{i=1}^{N_{\mathbf{T}}} (\mathbf{t}_i - \hat{\mu}_{\mathbf{T}})(\mathbf{t}_i - \hat{\mu}_{\mathbf{T}})^{\mathrm{T}} + \frac{N_{\mathbf{B}} N_{\mathbf{T}}}{N} (\hat{\mu}_{\mathbf{B}} - \hat{\mu}_{\mathbf{T}})^{\mathrm{T}} (\hat{\mu}_{\mathbf{B}} - \hat{\mu}_{\mathbf{T}}) \right| =$$

$$\begin{vmatrix} 1 & \sqrt{\frac{N_{\mathbf{B}} N_{\mathbf{T}}}{N}} (\hat{\mu}_{\mathbf{B}} - \hat{\mu}_{\mathbf{T}}) \\ \sqrt{\frac{N_{\mathbf{B}} N_{\mathbf{T}}}{N}} (\hat{\mu}_{\mathbf{B}} - \hat{\mu}_{\mathbf{T}}) & \sum_{i=1}^{N_{\mathbf{B}}} (\mathbf{b}_i - \hat{\mu}_{\mathbf{B}})(\mathbf{b}_i - \hat{\mu}_{\mathbf{B}})^{\mathrm{T}} + \sum_{i=1}^{N_{\mathbf{T}}} (\mathbf{t}_i - \hat{\mu}_{\mathbf{T}})(\mathbf{t}_i - \hat{\mu}_{\mathbf{T}})^{\mathrm{T}} \end{vmatrix}$$

The interchange of one row and one column leaves the determinate unchanged. So, using the identity again gives

$$\frac{L(\hat{\omega})}{L(\hat{\Omega})} = \left[\frac{1}{1 + \sqrt{\frac{N_{\mathbf{B}} N_{\mathbf{T}}}{N}} (\hat{\mu}_{\mathbf{B}} - \hat{\mu}_{\mathbf{T}})^{\mathrm{T}} \mathbf{A}^{-1} \sqrt{\frac{N_{\mathbf{B}} N_{\mathbf{T}}}{N}} (\hat{\mu}_{\mathbf{B}} - \hat{\mu}_{\mathbf{T}})} \right]^{\frac{N}{2}}$$

, where

$$\mathbf{A} = \sum_{i=1}^{N_{\mathbf{B}}} (\mathbf{b}_i - \hat{\mu}_{\mathbf{B}})(\mathbf{b}_i - \hat{\mu}_{\mathbf{B}})^{\mathrm{T}} + \sum_{i=1}^{N_{\mathbf{T}}} (\mathbf{t}_i - \hat{\mu}_{\mathbf{T}})(\mathbf{t}_i - \hat{\mu}_{\mathbf{T}})^{\mathrm{T}} .$$

The inequality $\frac{L(\varpi)}{L(\Omega)} \le k$ is equivalent to $\frac{N_{\mathbf{B}} N_{\mathbf{T}}}{N} (\hat{\mu}_{\mathbf{B}} - \hat{\mu}_{\mathbf{T}})^{\mathrm{T}} \mathbf{S}^{-1} (\hat{\mu}_{\mathbf{B}} - \hat{\mu}_{\mathbf{T}}) \ge \left(k^{\frac{-2}{N}} - 1 \right)(N - 2)$, where $\mathbf{S} = \mathbf{A} / (N - 2)$ is the sample covariance. Rewriting with the observed sample $\mathbf{X}$, H_0 is rejected if

$$d(\mathbf{X}) = \frac{N_{\mathbf{B}} N_{\mathbf{T}}}{N} (\mathbf{X}\mathbf{M}_{\mathbf{B}} - \mathbf{X}\mathbf{M}_{\mathbf{T}})^{\mathrm{T}} \mathbf{S}^{-1} (\mathbf{X}\mathbf{M}_{\mathbf{B}} - \mathbf{X}\mathbf{M}_{\mathbf{T}}) \ge \left(k^{\frac{-2}{N}} - 1 \right)(N - 2)$$

, where

$$\mathbf{S} = \frac{1}{N-2} \left[(\mathbf{B} - \mathbf{X}\mathbf{M}_{\mathbf{B}})(\mathbf{B} - \mathbf{X}\mathbf{M}_{\mathbf{B}})^{\mathrm{T}} + (\mathbf{T} - \mathbf{X}\mathbf{M}_{\mathbf{T}})(\mathbf{T} - \mathbf{X}\mathbf{M}_{\mathbf{T}})^{\mathrm{T}} \right]$$

The number k is selected so that the test has the desired level of significance. That level of significance is the probability of the inequality when H_0 is true, written $P\left(d(\mathbf{X}) > \left(k^{\frac{-2}{N}} - 1 \right)(N - 2) \mid H_0 \text{ true} \right)$. This is found by finding the distribution of $d(\mathbf{X})$.

3. DISTRIBUTION

The distributions of the background and target pixels were assumed to be $N(\mu_{\mathbf{B}}, \Sigma)$ and $N(\mu_{\mathbf{T}}, \Sigma)$. It can be easily shown that their sample means, $\hat{\mu}_{\mathbf{B}}$ and $\hat{\mu}_{\mathbf{T}}$, are therefore normally distributed according to $N(\mu_{\mathbf{B}}, \frac{1}{N_{\mathbf{B}}}\Sigma)$ and $N(\mu_{\mathbf{T}}, \frac{1}{N_{\mathbf{T}}}\Sigma)$ respectively. Another consequence is that $\delta = \left(\sqrt{\frac{N_{\mathbf{B}}N_{\mathbf{T}}}{N}}\right)(\hat{\mu}_{\mathbf{B}} - \hat{\mu}_{\mathbf{T}})$ is normally distributed. The terms $\mu_{\mathbf{B}}$ and $\mu_{\mathbf{T}}$ are the true means of the distributions of background and target. Under H_0, the distribution of δ is $N(0, \Sigma)$.

The term $\mathbf{S}$ was derived from $\mathbf{A} = \sum_{i=1}^{N_{\mathbf{B}}}(\mathbf{b}_i - \hat{\mu}_{\mathbf{B}})(\mathbf{b}_i - \hat{\mu}_{\mathbf{B}})^{\mathrm{T}} + \sum_{i=1}^{N_{\mathbf{T}}}(\mathbf{t}_i - \hat{\mu}_{\mathbf{T}})(\mathbf{t}_i - \hat{\mu}_{\mathbf{T}})^{\mathrm{T}}$. $\mathbf{A}$ has a Wishart distribution and therefore, so does $(N-2)\mathbf{S}$. That is $(N-2)\mathbf{S}$ can be written as $\sum_{i=1,2}\mathbf{Z}_i\mathbf{Z}_i^{\mathrm{T}}$, where each $\mathbf{Z}_i$ is independent with distribution $N(0,\Sigma)$. Under these conditions, it can be shown that $F = \left(\frac{N-p-1}{p(N-2)}\right)d(\mathbf{X})$ is distributed as a noncentral F with p and $N-p-1$ degrees of freedom and $\delta^{\mathrm{T}}\Sigma^{-1}\delta$ as the noncentrality parameter [5].

The probability of a false alarm (Pfa) is equal to the probability of rejecting H_0 when H_0 is true. The noncentrality parameter becomes zero under H_0. If the desired Pfa is α, then

$$\alpha = P\left(\frac{N_{\mathbf{B}}N_{\mathbf{T}}}{N}(\hat{\mu}_{\mathbf{B}} - \hat{\mu}_{\mathbf{T}})^{\mathrm{T}}\mathbf{S}^{-1}(\hat{\mu}_{\mathbf{B}} - \hat{\mu}_{\mathbf{T}}) > \left(k^{\frac{-2}{N}} - 1\right)(N-2) \,\Big|\, H_0 \text{ true}\right)$$

$$\alpha = P\left(\left(\frac{N-p-1}{p(N-2)}\right)\frac{N_{\mathbf{B}}N_{\mathbf{T}}}{N}(\mu_{\mathbf{B}} - \mu_{\mathbf{T}})^{\mathrm{T}}\mathbf{S}^{-1}(\mu_{\mathbf{B}} - \mu_{\mathbf{T}}) \geq \left(\frac{N-p-1}{p(N-2)}\right)\left(k^{\frac{-2}{N}} - 1\right)(N-2) \,\Big|\, H_0 \text{ true}\right)$$

$$= 1 - F_{0,\,p,\,N-p-1}\left(\left(\frac{N-p-1}{p}\right)\left(k^{\frac{-2}{N}} - 1\right)\right)$$

. Thus, k can be solved for explicitly.

The probability of detection (Pd) is equal to the probability of rejecting H_0 when H_1 is true. This is calculated as

$$\beta = P\left(\left(\frac{N-p-1}{p(N-2)}\right)\frac{N_{\mathbf{B}}N_{\mathbf{T}}}{N}(\mu_{\mathbf{B}} - \mu_{\mathbf{T}})^{\mathrm{T}}\mathbf{S}^{-1}(\mu_{\mathbf{B}} - \mu_{\mathbf{T}}) \geq \left(\frac{N-p-1}{p(N-2)}\right)\left(k^{\frac{-2}{N}} - 1\right)(N-2) \,\Big|\, H_1 \text{ true}\right)$$

$$= 1 - F_{\delta;\,p,\,N-p-1}\left(\left(\frac{N-p-1}{p(N-2)}\right)k^{\frac{-2}{N}} - 1\right).$$

Since the noncentrality parameter is nonzero under H_1, computation of the Pd requires knowledge of the target and background distributions.

4. EXPERIMENTAL RESULTS

The algorithm was applied to six-band multispectral data obtained from the Coastal Systems Station, Dahlgren Division Naval Surface Warfare Center, Panama City Florida. The filter wheel used six different spectral bands.

Figure 3 shows one band from a grass field, and figure 4 shows the strongest detection results from the grass scene. Figure 5 shows one band from a beach scene, and figure 6 shows the strongest detection results from the beach scene. The background region for both scenes was selected to be a 31 x 31 pixel mask surrounding a 5 pixel diameter target mask. The squares in the detection images indicate correctly identified targets. As can be seen in the detection images, the algorithm correctly identified the majority of targets with few alarms.

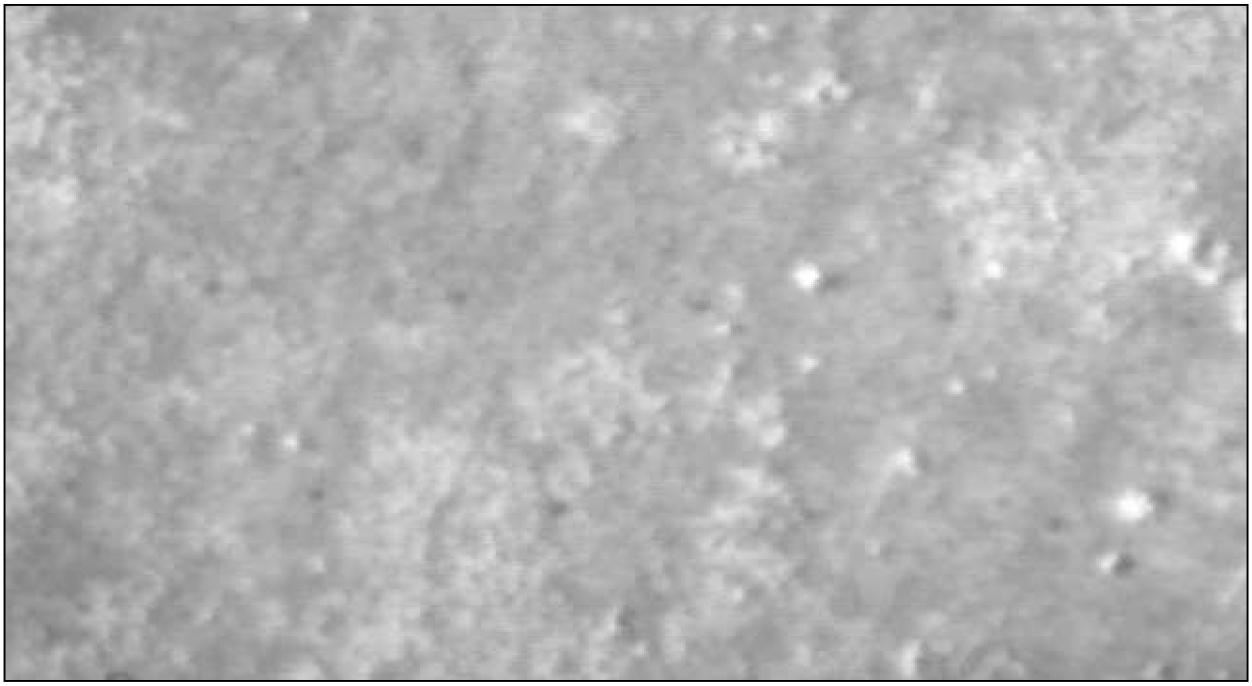

Figure 4: Band number six of a six band set showing a grassy area with dark targets.

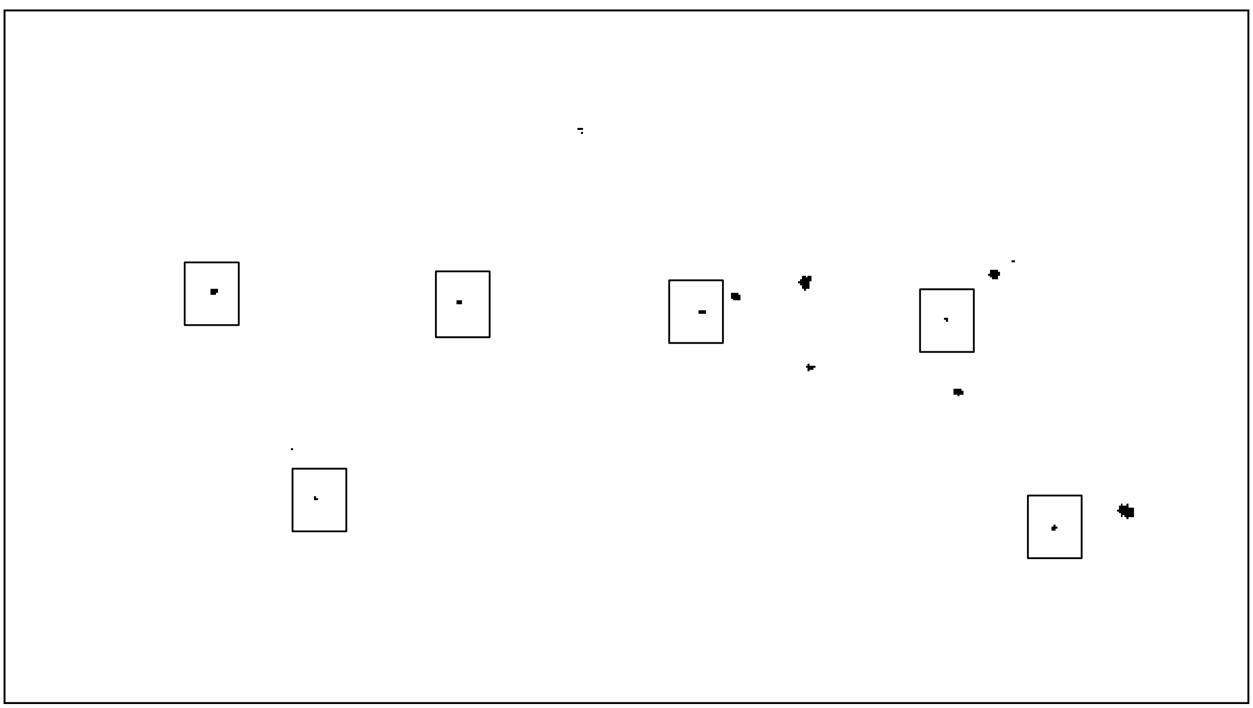

Figure 3. The highest confidence detections given by the algorithm on the grassy area.

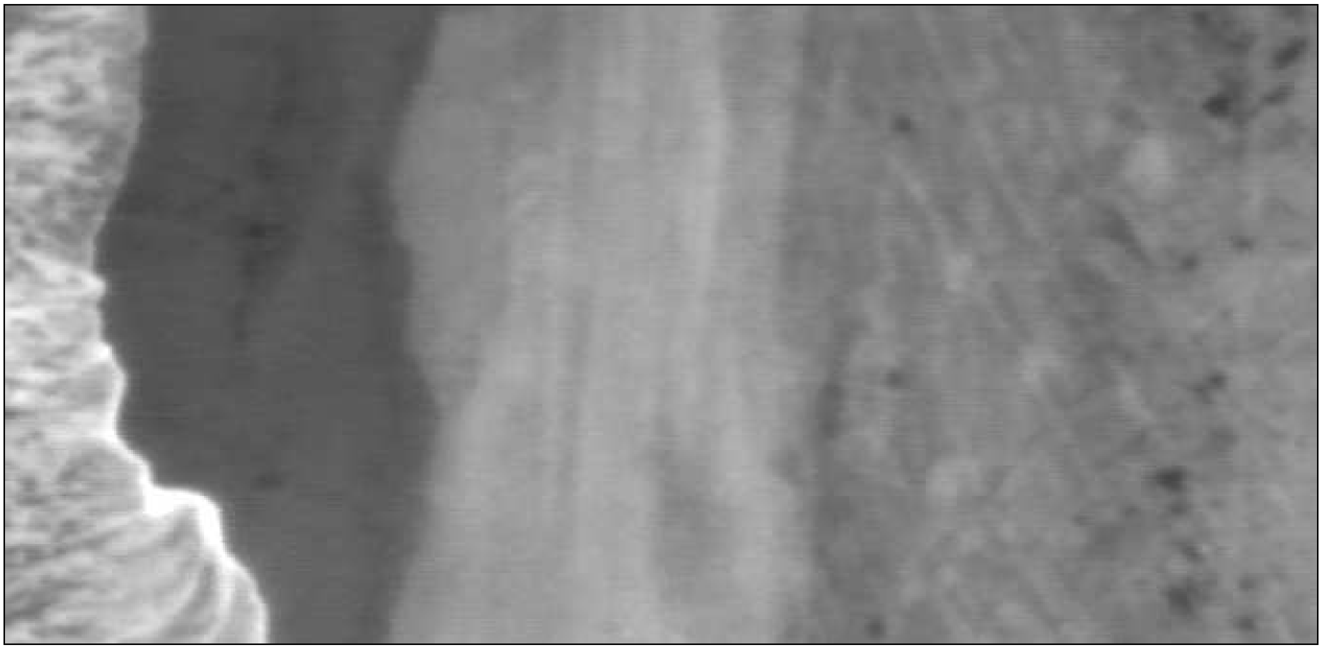

Figure 5. Band 3 of a six band set collected over a beach area

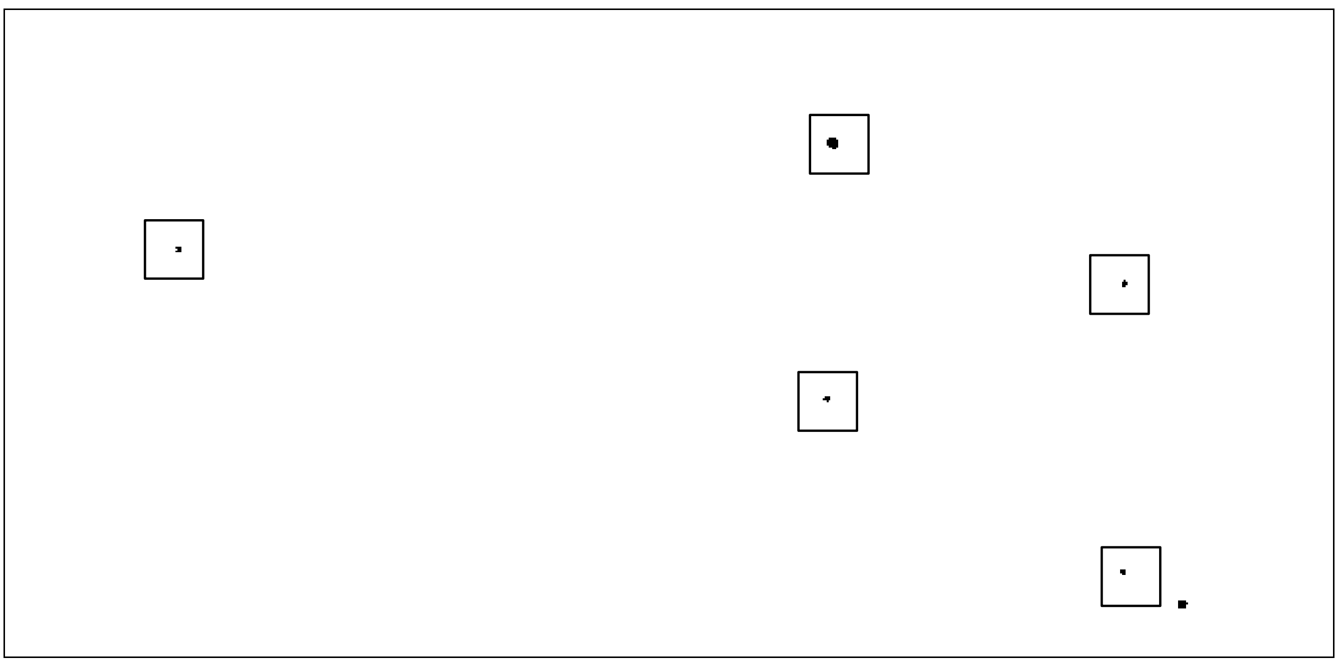

Figure 6. The highest confidence detections given by the algorithm on the beach area.

5. CONCLUSION

The CFAR algorithm presented detects targets at a specified false alarm rate. Actual Pd and Pfa results depend on how well the data match the model assumptions. The test statistic is derived from a ratio criterion which provides an optimal test of the hypotheses. The algorithm adapts to changing environmental conditions and target signatures by estimating characteristics based on the observed data. This includes cases where the targets may be lighter or darker than the surrounding background.

Further, there is no assumption regarding which bands or how many bands contribute to the spectral differences between the target and the background.

ACKNOWLEDGMENTS

The authors would like to thank Mr. Richard Miller of the NSWCDD-Coastal Systems Station USMC project office for providing the data presented in this paper.

REFERENCES

1. R. Muise, J.A. Wright, and Q.A. Holmes, "Coastal Mine Detection Using the COBRA Multispectral Sensor", *Proceedings of SPIE* **2765**, April 1996.

2. B.R. Hunt and T.M. Cannon, "Non-stationary Assumptions for Gaussian Models of Images", *IEEE Transactions on Systems, Man and Cybernetics*, **SMC-6** 1976.

3. I.S. Reed and Xiaoli Yu, "Adaptive Multiple-Band CFAR Detection of an Optical Pattern with Unknown Spectral Distribution", *IEEE Transactions of Acoustics, Speech, and Signal Processing*, **38** No. 10 Oct., 1990.

4. R.V. Hogg and E.A. Tanis, *Probability and Statistical Inference*, Macmillan Publishing Co. Inc., New York, 1983.

5. D.F. Morrison, *Mutivariate Statistical Methods*, McGraw-Hill, New York, 1976.