Point Target Detection in Spatially Varying Clutter

Srishan Sridhar and Glenn Healey
Department of Electrical and Computer Engineering
University of California
Irvine, CA 92717

Abstract

In this paper we develop and analyze high-speed algorithms for the detection of point targets in infrared (IR) images with spatially varying clutter. Current target detection systems are effective in detecting bright targets in a uniform sky, but in areas of strong clutter are either unable to detect targets reliably or are limited by high false alarm rates. We assume that target and sensor models are available. Clutter is considered to be poorly characterized and spatially varying. Target detection algorithms are based on filtering to enhance the target signal relative to the background, followed by an adaptive threshold. Statistical analysis of the algorithms is provided to quantify algorithm performance. Our system implements a spatially adaptive algorithm that maximizes probability of target detection while maintaining a fixed false alarm rate. The algorithms are robust in the presence of spatially varying clutter. We include experimental results to illustrate this.

1 Introduction

Although target detection algorithms have been steadily improving [2], target detection systems are still only partially successful. Current systems are often unable to detect targets reliably in the presence of other bright objects in the scene. In addition, many systems are unable to adapt their behavior to perform robustly in the rapidly changing unconstrained environments that are frequently encountered.

Under ideal circumstances, a point target will appear as a bright spot in an IR image against a uniform background. For this case, a single global threshold will be sufficient for target detection. More generally, the image will also contain bright clutter regions due to scene events such as brightly illuminated terrain or sunlit clouds. In this situation, a single threshold will not be sufficient to discriminate targets from the equally bright clutter regions.

Since targets and clutter have different spatial frequency characteristics, a spatial filter could be designed to suppress clutter and detect target signals. Ideally, such a filter would produce a strong response only in the presence of a target and a subsequent thresholding of the filtered image would be adequate for target detection. To design such a filter requires knowledge about the characteristics of targets and clutter. From the point spread function of our sensor, we can predict the signal that will be produced by a point target. This signal is often modeled in one dimension, for example, as a gaussian or raised cosine [4]. Unfortunately, accurate models for clutter are often not available for real applications. Furthermore, clutter properties frequently change with location in the image.

2 The Target Detection Problem

A standard technique for detecting a specified target signal in an image is to derive a matched filter for the target signal. Since the target signal is known, we could design an appropriate matched filter for point target detection if we knew the power spectrum of the background. The power spectrum of the background consists of a clutter term, a white-noise term, and cross-spectral terms between clutter and white-noise [5]. Except at high frequencies, the clutter term typically dominates the terms involving white noise. Thus, the biggest difficulty in deriving a matched filter for point target detection involves choosing a model for the power spectrum of the clutter. The choice of such a model must be made with care because the performance of a matched filter will degrade significantly when it is applied to images whose background spectrum is different from the background spectrum used to derive the matched filter [8].

Several studies of infrared data have led to clutter models that are reasonable for certain classes of scenes [1] [3]. The most serious disadvantage of using such

0-8186-2840-5/92 $03.00

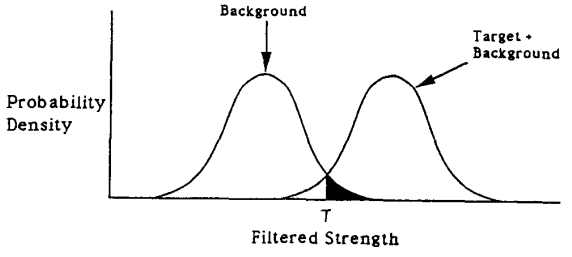

Figure 1: Filtered distributions

models is that they force the choice of a matched filter that is tuned to some average properties of the anticipated clutter. A matched filter derived from such models cannot respond to spatially varying clutter. Thus, for example, it is not reasonable to expect good performance from such a filter in both clear sky where the background is dominated by white noise and in strong clutter where the background power spectrum is primarily concentrated at lower frequencies.

An alternative approach to using a matched filter for target detection is to use a least-mean-square (LMS) filter [8]. An LMS filter separates an input image into a target component, a background component and a residual component with only the target component being passed by the filter. The background component is all of the input that can be represented using the background model. Input that is not assigned to background that can be fit by the target model is assigned to the target component. The residual component is remaining input. For any input image, each of the three components will contain some amount of actual target signal and background. Therefore, the performance of the filter is critically dependent on how well the image background can be represented by the assumed background model.

If filtering is effective, the probability distribution of target plus background in the filtered image will be well separated from the distribution of background (figure 1). A threshold T can then be applied to the filtered image to detect targets. The false alarm rate for a given choice of T is indicated by the shaded area in figure 1 .

3 Analyzing Matched and LMS Filters

We quantify the performance of matched and LMS filters derived using various background assumptions as a function of the properties of the actual background. In this way, we examine analytically not only how well each filter performs on images for which it is well suited, but also how the performance of the filter degrades as the assumed background model becomes inaccurate. Our analysis allows us to determine the consequences of using different filters and background assumptions for various classes of images.

3.1 Background and Target Models

We consider backgrounds formed by passing white gaussian noise with mean μ and variance σ^2 through a butterworth filter with transfer function

$$H(\omega_x, \omega_y) = \frac{1}{\sqrt{1 + (\omega_x^2 + \omega_y^2)/c^2}} \qquad (1)$$

where ω_x and ω_y are the 2-D spatial frequency variables, and c is a constant. We have studied backgrounds with values of c ranging from 10 to 1000. The output of the butterworth filter satisfies the isotropic power spectral density

$$S(\omega_x, \omega_y) = \frac{\sigma^2}{1 + (\omega_x^2 + \omega_y^2)/c^2} + \mu^2 \delta(\omega_x, \omega_y) \qquad (2)$$

Smaller values of c correspond to backgrounds which are almost dc. Large values of c correspond to backgrounds which are almost white gaussian noise. Figure 2 shows 1-D plots of the power spectra as functions of c when $\mu = 0$ and $\sigma^2 = 1$. The butterworth filter is such that the mean of the resulting filtered image equals the mean of the input white gaussian noise.

The target signal is modelled as a discrete raised-cosine pulse in 2-D.

$$T(R, i, j) = \begin{cases} \left[1 + \cos \frac{\pi}{R} \sqrt{i^2 + j^2}\right] & \text{if } i^2 + j^2 < R^2 \\ 0 & \text{otherwise} \end{cases} \qquad (3)$$

where R is the radius in pixels. We assume a square support for our target to make use of the symmetry of the raised-cosine. Therefore, the size of the target is $2R - 1 \times 2R - 1$ pixels. We will consider in detail the case $R = 4$. We assume that targets and background are additive. Hence we consider backgrounds described by (2) with scaled target signals added at different positions.

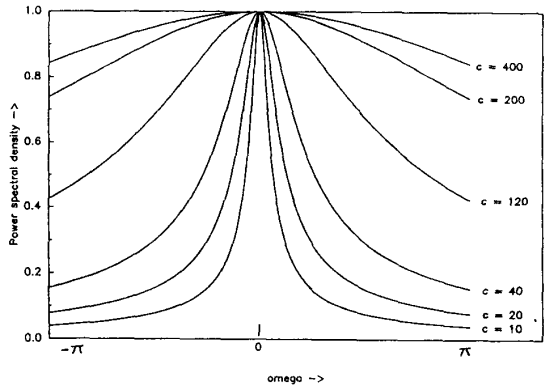

Figure 2: 1-D plots of power spectra.

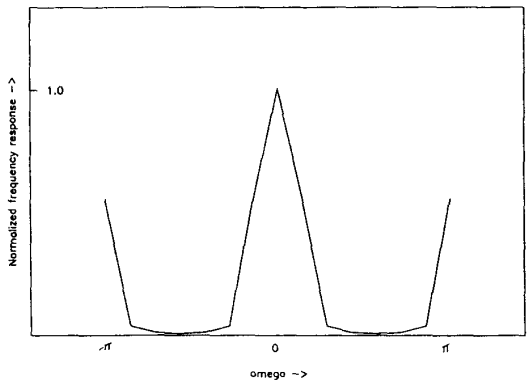

Figure 3: Matched filter frequency response.

3.2 Matched filters

Discrete matched filters were derived for backgrounds with c varying between 10 and 1000 and containing targets that are scaled versions of $T(R,i,j)$. Using a vector space formulation, let r denote the received digital image vector containing target signal and additive background. Let t and n represent the target signal vector and background vector respectively. The elements of t are obtained by stacking the rows of $T(R,i,j)$ in one column. If the target is $2R-1 \times 2R-1$ pixels in size, then r,t and n are each $2R-1 \times 2R-1$ element column vectors. Thus $r = t+n$ in vector space formulation. Note that the mean of vector n is $\mu\vec{1}$, where $\vec{1} = [1\ 1\ \ldots]^T$, since the butterworth filter does not alter the mean of its input. If m denotes the matched filter in vector notation, then [6]

$$m = K_n^{-1}t \tag{4}$$

where $K_n = cov[n] = E[(n-\mu\vec{1})(n-\mu\vec{1})^T]$ and denotes the covariance matrix of the background for which the matched filter is derived, i.e. the assumed background. From (2), K_n is a function of σ^2 and c alone. Moreover it is directly proportional to σ^2. Hence we can write $K_n = \sigma^2\widehat{K_n}$ where $\widehat{K_n}$ is a function of c alone. For white noise $K_n = \sigma^2 I$. When $R = 4$, t and m are 49 element vectors and hence K_n is a 49×49 matrix. Figure 3 shows the normalized frequency response of the matched filter for $c = 400$ when $\sigma^2 = 4$ and $R = 4$.

When the matched filter of (4) is applied to a background whose covariance matrix is $K_b \neq K_n$, the resulting signal-to-noise ratio (filtered target power to filtered background power) is

$$\text{SNR} = \frac{(\mathbf{m}^T\mathbf{t})^2}{\mathbf{m}^T\mathbf{K_b}\mathbf{m}} = \frac{(\mathbf{t}^T\mathbf{K_n^{-1}}\mathbf{t})^2}{\mathbf{t}^T\mathbf{K_n^{-1}}\mathbf{K_b}\mathbf{K_n^{-1}}\mathbf{t}} \tag{5}$$

When $K_b = K_n$ we get the SNR value of $t^T K_n^{-1} t$.

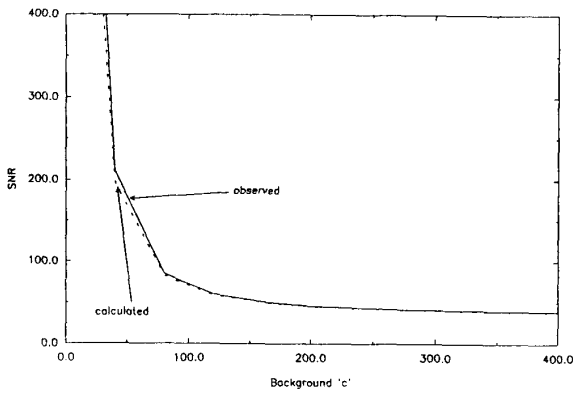

Figure 4: SNR when applied to different backgrounds.

We have studied the degradation of the matched filter when it is applied to backgrounds for which it is not matched, i.e. $K_b \neq K_n$. The SNR is a function of R, σ^2, and c for both the assumed and actual backgrounds. We made theoretical calculations of the SNR for the cases when the matched filters were applied to backgrounds to which they were not matched and followed it up with experiments. Figure 4 shows how the SNR degrades when the filter matched to $c = 10$ is applied to different backgrounds when $R = 4$ and $\sigma^2 = 4$.

From our background model, the component of the filtered image due to background is gaussian. Hence the filtered background component is described by its

mean and variance. The mean of the filtered background is

$$\mu_{filtered} = E[m^T n] = m^T \mu \vec{1}$$
$$= t^T K_n^{-1} \mu \vec{1} = (\mu/\sigma^2) t^T \widehat{K_n}^{-1} \vec{1} \quad (6)$$

We note that $\mu_{filtered}$ is proportional to μ and is a function of the target signal and c of the assumed background $\left(\text{in } \widehat{K_n}^{-1}\right)$.

The variance of the filtered background is given by

$$\sigma^2_{matched} = E[(m^T n - m^T \mu \vec{1})^2]$$
$$= E[m^T (n - \mu \vec{1})(n - \mu \vec{1})^T m] \quad (7)$$

If we assume that the filter is applied to the background to which it is matched, then

$$\sigma^2_{matched} = m^T K_n m = t^T K_n^{-1} t = \frac{1}{\sigma^2} t^T \widehat{K_n}^{-1} t \quad (8)$$

When the filter is not matched to the background we get

$$\sigma^2_{unmatched} = t^T K_n^{-1} K_b K_n^{-1} t \quad (9)$$

where K_b is the covariance matrix of the actual background. Tables showing calculated and observed values of the filtered background mean and variance for the case $R = 4$, $\mu = 10$, and $\sigma^2 = 10$ can be found in [7].

3.3 LMS Filters

In using an LMS filter, a model of the target-plus-background is fit by least-squares at every point in the sensed image. The target model is the product of a scale factor A and the shape function $T(R, i, j)$. The background model is an arbitrary function represented by a few terms of a truncated Maclaurin series. This target-plus-background model is fit over a local neighborhood of the pixel to obtain an estimate of the target strength A at that pixel. Figure 5 shows the sample labeling and coordinate system for the case $R = 4$. For this case, the raised-cosine target is zero valued at the corners and only the pixels marked by a cross are taken into account in the least squares fit. With these conventions, the first order target-plus-background model for K samples ($K = 45$ in figure 5) can be written in element form as

$$g_j = At_j + B^{0,0} + B^{1,0} x_j + B^{0,1} y_j$$
$$j = 1, 2, \ldots, K \quad (10)$$

where g_j denotes the jth sample to which the model is fit, t_j is the jth sample of the target, and x_j, y_j

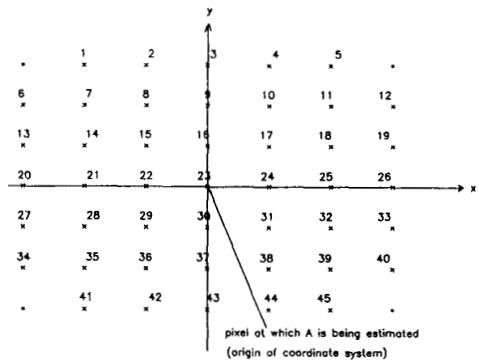

Figure 5: Samples to which the model is fit.

represent the x and y coordinates of the sample point with respect to the origin of the coordinate system (figure 5). A, $B^{0,0}$, $B^{1,0}$, and $B^{0,1}$ are adjusted to achieve the least-squares fit. The first order LMS filter makes these adjustments spontaneously at every pixel in the image.

If r_j is the actual observed sample in the received image, the error due to the model is $r_j - g_j$. Since the target signal is symmetric, the estimate of A at a pixel in the image (first order model) is given by [7]

$$\hat{A} = \sum_{j=1}^{K} \left[(Kt_j - \sum_{p=1}^{K} t_p) r_j \right] = \sum_{j=1}^{K} l_j r_j \quad (11)$$

The right hand side of (11) can be interpreted as a 2-D convolution sum in which l_j are the first order LMS filter weights. A detailed derivation of the first and second order LMS filter weights for the target signal of (3) with R=4 can be found in [7]. Note that for a given background model, the LMS filter weights depend only on R. Figure 6 shows the normalized frequency response of the first order LMS filter in one dimension. Notice how the frequency response goes to zero for the dc point at the center of the figure.

Since $r_j = t_j + n_j$, the output of the first order LMS filter applied to r can be written as

$$\hat{A} = \left[\sum_{j=1}^{K} (Kt_j - \sum_{p=1}^{K} t_p) t_j \right] + \left[\sum_{j=1}^{K} (Kt_j - \sum_{p=1}^{K} t_p) n_j \right] \quad (12)$$

Therefore, it consists of a target and a background component. It is easily verified that $\sum_1^K l_j = 0$ and thus the background component has zero mean. Hence

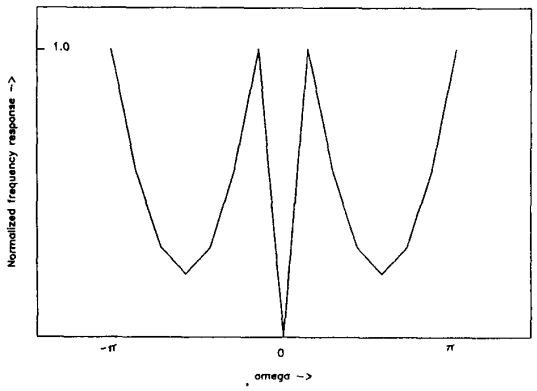

Figure 6: LMS filter frequency response.

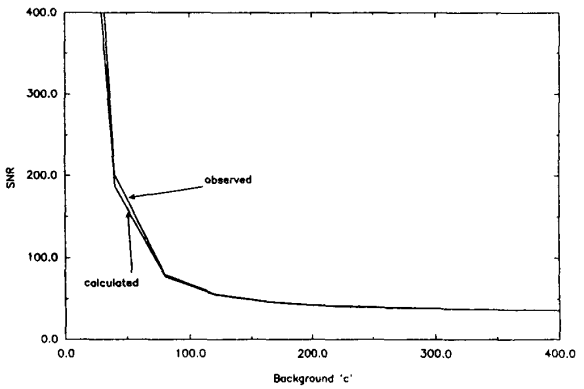

Figure 7: SNR variation as a function of c.

we can write the output SNR as

$$\text{SNR} = \frac{\left[\sum_{j=1}^{K}(Kt_j - \sum_{p=1}^{K} t_p)\right]^2}{\sum_{j=1}^{K}\sum_{m=1}^{K}(Kt_j - \sum_{p=1}^{K} t_p)(Kt_m - \sum_{q=1}^{K} t_q)P_{jm}} \quad (13)$$

where $P_{jm} = E[n_j n_m]$. If the background is zero mean unit variance white noise, $P_{jm} = \delta_{jm}$. P_{jm} is a function of μ, σ^2 and c. Figure 7 shows the SNR variation when the first order LMS filter is applied to different backgrounds for the case when $R = 4$, $\mu = 10$, and $\sigma^2 = 4$. It is seen that the performance of the LMS filter degrades considerably for higher values of c. This is due to the fact that LMS filters pass a large amount of high frequency background as illustrated in figure 6.

As an example, suppose the background is zero mean unit variance white gaussian noise. Then, $P_{jm} = \delta_{jm}$ and since $K = 45$,

$$\frac{\text{Matched filter SNR}}{\text{LMS filter SNR}} = \frac{t^T t}{t^T t - (\sum_{j=1}^{K} t_j)^2/K} = 2.35 \quad (14)$$

where as before $t = [t_1\ t_2 \cdots t_K]$ is the target vector. Thus the matched filter performs better than the first order LMS filter for white noise by a factor of 2.35 for the chosen target model.

Let l denote the vector form of the LMS filter. The output of the LMS filter due to background is $l^T n$. This is gaussian as can be readily observed. Hence, it is described by its mean and variance. Since the LMS filters (both first and second order) reject dc, the filtered background has zero mean. The variance of the LMS filtered background is

$$\sigma^2_{LMS} = E[(l^T n)^2] = l^T K_n l = \sigma^2 l^T \widehat{K_n} l \quad (15)$$

Tables showing calculated and observed values of LMS filtered background variance for the case when $R = 4$, $\mu = 10$, and σ^2 can be found in [7].

4 Thresholding

4.1 Global Thresholding

The purpose of filtering is to separate as much as possible the target-plus-background distribution from the background distribution in the filtered image. Typically, these distributions will overlap after filtering so that a single threshold will not be able to identify all targets without false alarms. In this section, we examine two measures for the separability of filtered distributions. The MSV is defined as the absolute difference between the means of the filtered background and filtered target-plus-background distributions divided by the standard deviation of the filtered background. We note that the filtered target-plus-background distribution has the same standard deviation as the filtered background. Another useful measure of separability is the probability of target detection for a threshold chosen to achieve a given false alarm rate. The probability of detection for a threshold T is equal to the area of the filtered target-plus-background distribution between T and infinity (figure 1).

In section 3.2, we derived expressions for the mean $\mu_{filtered}$ and variance of background processed by a matched filter. The distribution of filtered target-plus-background has the same variance as the filtered

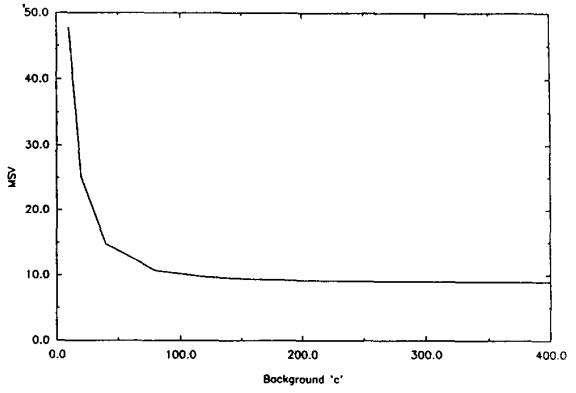

Figure 8: MSV for the matched filter case.

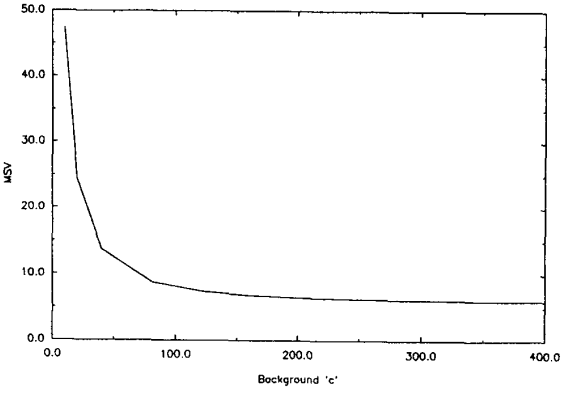

Figure 9: MSV for the LMS case.

background and a mean given by $m^T t + \mu_{filtered}$. The mean separation for matched filtering is given by $m^T t$ and the variance of the filtered background is given by (7). A plot of MSV as a function of c when filters are applied to backgrounds for which they are matched is shown in figure 8. Here we have assumed $R = 4$, $\mu = 10$, and $\sigma^2 = 4$.

In section 3.3, we showed that background processed by an LMS filter has zero mean and a variance given by (15). Thus, the mean separation for an LMS filter is $l^T t$. A plot of MSV as a function of c for a first order LMS filter is shown in figure 9. As seen from the figure, the LMS filter produces good separation for small c but poorer separation as c becomes large and the background approaches white noise.

In many applications, we desire to maximize the probability of target detection while attaining a prescribed false alarm rate. From figure 1, we see that a threshold that achieves a given false alarm rate can be determined if the distribution of filtered background can be estimated. Such estimates can be made either from the filtered image or by applying the analysis of section 3 to estimates of the actual background properties.

4.2 Adaptive Filtering and Thresholding

Thus far we have characterized each background using a stationary model having a power spectral density given by (2). However, there often is considerable spatial variation in the statistical properties of clutter. If we assume that the clutter is stationary over regions of the image with a local power spectral density given by (2), then we can partition a background image into different regions each with a unique value of c. We partition the image by dividing it into rectangular regions and estimating a value of c for each region to characterize the local power spectral density. Appropriate filters and thresholds can then be applied to each region thereby improving overall system performance.

The response of a matched filter is highly dependent on the c of the background. We have seen that matched filter performance degrades considerably when applied to backgrounds other than it is matched for. Thus, having subdivided the image into different regions, we can apply the filter that is matched to the estimated c of each region. In this way, we can enhance the SNR locally over the image rather than applying a single matched filter using globally averaged statistics.

The LMS filter automatically fits the received image to the target-plus-background model. The performance of the LMS filter depends on how well the deterministic target-plus-background model fits the background. Even though one LMS filter is sufficient for the whole image, partitioning an image into different regions assists in determining local threshold values based on the estimated c of each region.

5 Experimental results

We applied our adaptive filtering and thresholding techniques to 480 x 640 pixel infrared images of point targets in spatially varying clutter. We divided each image into 64 rectangular regions of 60 x 80 pixels and estimated the parameters c and σ^2 of our background model for each region. Each region of the image was filtered using the appropriate matched filter. We also filtered the image using a first order LMS filter. From the estimates of c and σ^2, we used the analysis of section 3 to compute 2% false alarm thresholds for

each region. This adaptive thresholding led to low overall false alarm rates and succeeded in providing a high probability of detection.

We present some images demonstrating our approach. Figure 10 is an IR image containing 6 targets embedded in spatially varying clutter. Figures 11 and 12 show the LMS filtering and adaptive matched filtering of figure 10. Figures 13 and 14 show the results of the adaptive thresholding procedure applied to the images of figures 11 and 12. The overall false alarm rate was lower than 2% and 5 of the 6 targets were detected in each case. We also thresholded the images of figures 11 and 12 using a single global threshold. Figures 15 and 16 show the results of this global thresholding. We obtained a very high false alarm rate and low probability of detection in each case. This demonstrates the ability of our adaptive thresholding procedure to reduce false alarm rate and improve probability of detection.

6 Conclusion

We have analyzed the properties of high speed target detection systems based on a filtering step for target enhancement and background suppression followed by a thresholding step for target detection. We have assumed a known target signature buried in a spatially varying background that is characterized by a parametric power spectral model. Using these models, we examined the usefulness of matched filters and LMS filters for target detection. Our analysis is useful for quantifying the performance of target detection systems and has been demonstrated by experiments with real and synthesized infrared images.

We have shown that in cases where image backgrounds have mostly low frequency content, the use of a single LMS filter is often sufficient for reliable target detection. LMS filters based on polynomial background models, however, pass significant high frequency background. In cases where high frequency backgrounds are possible, the use of matched filters tuned to the background is a better approach. Since matched filters depend on the exact power spectrum of the background, we have developed an adaptive matched filtering approach that is based on local spectral estimates and subsequent filter selection. For many classes of imagery this adaptive matched filtering approach is more effective than LMS filtering at the cost of additional computation.

Figure 10: IR image with targets.

Figure 11: LMS output.

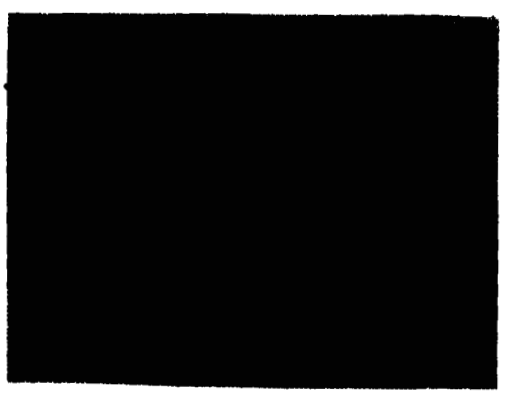

Figure 12: Adaptive matched filter output.

Figure 13: Thresholded LMS output.

Figure 14: Thresholded matched output.

Figure 15: Single thresholding of LMS output.

Figure 16: Single thresholding of matched output.

Acknowledgment

This research was supported by Loral Aeronutronic and the UC MICRO program.

References

[1] A. Ben-Yosef, K. Wilner, S. Simhony, and G. Feigin. Measurement and analysis of 2-d infrared natural background. *Applied Optics*, 24:2109, 1985.

[2] B. Bhanu. Automatic target recognition: State of the art survey. *IEEE Transactions on Aerospace and Electronic Systems*, 22(4), July 1986.

[3] A. Itakura, S. Tsutsumi, and T. Takagi. Statistical properties of the background noise for the atmospheric windows in the intermediate infrared region. *Infrared Physics*, 14(17), 1974.

[4] M. Longmire, A. Milton, and E. Takken. Simulation of mid-infrared clutter rejection. *Applied Optics*, 21(21):3819–3833, November 1982.

[5] P. Maybeck. *Stochastic Models, Estimation, and Control.* Academic Press, New York, 1979.

[6] W. Pratt. *Digital Image Processing.* Wiley, New York, 1978.

[7] S. Sridhar and G. Healey. Detecting point targets in spatially varying clutter. Technical Report ECE-92-08, UC Irvine, Aug 1992.

[8] E. Takken, D. Friedman, A. Milton, and R. Nitzberg. Least-mean-square spatial filter for infrared sensors. *Applied Optics*, 18(24):4210–4222, December 1979.