Performance evaluation of quadratic correlation filters for target detection and discrimination in infrared imagery

S. Richard F. Sims
U.S. Army Research, Development, and Engineering Command
Aviation and Missile Research, Development, and Engineering Center
AMSRD-AMR-SG-IP
Redstone Arsenal, Alabama 35898

Abhijit Mahalanobis, FELLOW SPIE
Lockheed Martin
Missiles and Fire Control Systems
Orlando, Florida 32801

Abstract. The detection and discrimination of targets in infrared imagery has been a challenging problem due to the variability of target and clutter (background) signatures. We discuss the application of a novel quadratic filtering method using missile seeker infrared closing sequences. Image filtering techniques are well suited for target detection applications, since they avoid the disadvantages of typical pixel-based detection schemes (such as segmentation and edge extraction). Another advantage is that the throughput complexity of the filtering approach, in the detection process, also does not vary with scene content. The performance of the proposed approach is assessed on several datasets, and the results are compared with that of previous linear filtering techniques. Since we can obtain the signature of some of the clutter ''in the field'' or during operation, we examine the impact of updating the filters to adapt to the clutter.
[DOI: 10.1117/1.1767195]

Subject terms: automatic target recognition; detection; discrimination; automatic target recognition performance analysis; correlation filters; quadratic correlation filters.

Paper TPR-013 received Aug. 21, 2003; revised manuscript received Feb. 13, 2004; accepted for publication Feb. 17, 2004.

1 Introduction

Over the past three decades, considerable work has been done in the area of detection and recognition of targets in infrared (IR) imagery. Popular methods for achieving this goal include statistical techniques based on feature classification,[1,2] model-based techniques,[3,4] and neural networks.[5–7] Generally speaking, the potential targets are segmented from the background; boundary edges and internal features are measured, and fitted to either statistical or structural models. These methods work well when there is a sufficient number of pixels on the targets and the contrast boundaries are well defined. In general, however, the performance of most segmentation-based methods suffers considerably when the targets exhibit poor contrast relative to the background, or when parts of the target are either obscured or in close proximity to clutter objects. In such cases, segmentation is a difficult task that generally induces significant error.

We focus on target detection techniques based on signal processing methods that do not require segmentation. The leading approach is based on the use of correlation filters.[8,9] Perhaps the earliest and most well known among these is the matched spatial filter (MSF).[10] Although the MSF is optimum for the detection of a single (deterministic) pattern in the presence of additive white Gaussian noise (AWGN), it is hardly suitable for robust detection and recognition of targets in real-world images. For our case, however, to add robustness and distortion tolerance to the detection process, we consider the data and the pattern as both being a stochastic process. This no longer fits the model in which the MSF is optimal and a different methodology is needed to define correlators for the detection problem. As a result, many approaches have been proposed for the design of filters for correlators that exhibit robust behavior in the presence of clutter and wide variations in the targets' signatures.

As outlined in Ref. 8, much work has been done on the design of linear filters for pattern recognition. In addition to locating the target in the input signal (or image), the filter design must accommodate the suppression of side lobes on the target of interest and false alarms. In this work, we evaluate a new type of quadratic correlation filter (QCF) that improves on linear correlation filters, but retains the advantage of shift invariance without requiring segmentation or registrations. The method reviewed here for the design of QCFs optimizes a performance metric, formulated as a Rayleigh quotient, leading to eigenspace solutions for QCFs.

Our focus is on comparing the performance of the maximum average correlation height (MACH) linear filters[11] and the QCFs in the detection of targets and in the subsequent discrimination based on the detections in infrared image ''closing'' sequences. We plot the detection metrics on the target and other objects versus range. The detection metrics used are a confidence measure for the QCF and the peak-to-side lobe ratio for the MACH. This type of plot or strip chart clearly shows the separation and in some cases nonseparation of the target versus clutter. In our closing sequence scenarios, which are from long to close range, which varies the resolution on the target, clutter is image

dependent, so that anything other than the object of interest encoded into the detection filter is deemed to be clutter. This clutter can include other targets, even of the same type if they are present in the same image. In addition, the clutter can also clearly include any scene artifacts, man-made or natural. As the infrared camera moves from long to close range, the aspect and perspective of the field of view onto the background changes, along with the range to target. In general, this scene background change and variation in pixel resolution on the target has the potential to drastically vary the target-like or clutter objects in the scene.

In the next sections, we discuss the theory and implementation of QCF filters, specifically the version that has been called the tuned basis function (TBF). We then describe the detection metrics, the data we used, and the performance quantification of the TBF versus the MACH filters.

2 Review of QCF Theory

Although the theory of linear correlation filters (including the MACH filter) is well known, there is less literature available on the design of QCFs. For ease of reference, we now review a technique for the design of QCFs. The issue of finding targets in background clutter, in this case, is a two-class pattern recognition problem. Consider an input training signal

$$\mathbf{x}=[x(0)\ \ x(1)\ \ \cdots\ \ x(N-1)]^T,$$

that can be either a target (class ω_1) or background (class ω_2). For now we assume that $\mathbf{x}$ is a purely real signal. The output of the quadratic filter is defined as

$$\mathbf{x}^T\mathbf{T}\mathbf{x}=\sum_{i=1}^{N}\sum_{j=1}^{N} t_{ij}x(i)x(j). \tag{1}$$

The coefficient matrix $\mathbf{T}=\{t_{ij}\}$ is square and real but otherwise unrestricted. It should be noted that since $\mathbf{T}$ is not restricted to be positive definite, the output can be either positive or negative. The idea is to determine $\mathbf{T}$ such that the output is positive and as large as possible when $\mathbf{x}\in\omega_1$, and negative or as small as possible when $\mathbf{x}\in\omega_2$. Assume that the form of the coefficient matrix for a quadratic filter is

$$\mathbf{T}=\sum_{i=1}^{N}\mathbf{q}_i\mathbf{q}_i^T, \tag{2}$$

where $\mathbf{q}_i$ form an orthonormal basis set. The output φ of the QCF in response to an input training vector $\mathbf{x}$ is given by

$$\varphi=\mathbf{x}^T\mathbf{T}\mathbf{x}. \tag{3}$$

The idea is to maximize the separation between the expected values of φ for the two classes while reducing an upper bound on the variance of φ. It can be shown that the separation between the expected value for the two classes is given by Eq. (4), where

$$E_1\{\varphi\}-E_2\{\varphi\}=\sum_{i=1}^{N}\mathbf{q}_i(\mathbf{R}_1-\mathbf{R}_2)\mathbf{q}_i^T, \tag{4}$$

where $\mathbf{R}_i$ is the correlation matrix of class ω_i. The maximization of this term alone, however, results in a solution that does not adequately control the spread of the class statistic. To avoid this, we also seek to minimize

$$E_1\{\varphi\}+E_2\{\varphi\}=\sum_{i=1}^{N}\mathbf{q}_i(\mathbf{R}_1+\mathbf{R}_2)\mathbf{q}_i^T. \tag{5}$$

It can be shown that including this term places bounds on the second-order statistics of φ, and helps to reduce the spread. To achieve our objectives, we seek $\mathbf{q}_i$ to maximize the ratio shown in Eq. (6).

$$J(\mathbf{q})=\frac{E_1\{\varphi\}-E_2\{\varphi\}}{E_1\{\varphi\}+E_2\{\varphi\}}=\frac{\Sigma_{i=1}^{N}\mathbf{q}_i(\mathbf{R}_1-\mathbf{R}_2)\mathbf{q}_i^T}{\Sigma_{i=1}^{N}\mathbf{q}_i(\mathbf{R}_1+\mathbf{R}_2)\mathbf{q}_i^T}. \tag{6}$$

By taking the derivative of the expression in Eq. (6) with respect to $\mathbf{q}_i$ and setting it equal to zero, we then, with some manipulations, have Eq. (7).

$$(\mathbf{R}_1+\mathbf{R}_2)^{-1}(\mathbf{R}_1-\mathbf{R}_2)\mathbf{q}_i=\lambda_i\mathbf{q}_i. \tag{7}$$

Thus $\mathbf{q}_i$ is an eigenvector of the matrix $(\mathbf{R}_1+\mathbf{R}_2)^{-1}(\mathbf{R}_1-\mathbf{R}_2)$ with eigenvalue λ_i. The dominant eigenvector of course maximizes the ratio for the choice of any one vector. However, by imposing the constraint that $\mathbf{q}_i$ must form an orthonormal set, it is seen that the optimum choice for N vectors is to select N eigenvectors of $(\mathbf{R}_1+\mathbf{R}_2)^{-1}(\mathbf{R}_1-\mathbf{R}_2)$ in order of decreasing eigenvalues.

3 Efficient Implementation of QCFs

Consider now the problem of processing an entire 2-D image using the quadratic filter. Let us assume that the test input vector $\mathbf{z}$ is obtained by lexicographically reordering the data in a relatively small window of row and column size $r\times c$ within a much larger image of size $R\times C$. (Typical sizes of real-world images may be 256×256 pixels (or larger), whereas the window size for $\mathbf{x}$ may be 20×20, chosen to be approximately the size of the targets of interest.) The complexity of the direct computation necessary to evaluate the expression in Eq. (1) at all overlapping $r\times c$ subregions of the image is rather large. We therefore seek an efficient architecture for obtaining the full correlation of the input scene and the quadratic filter.

Denoting the dimensionality of the space by $d=r\cdot c$, let us assume that the $d\times d$ coefficient matrix $\mathbf{T}$ of rank N can be expressed as $\mathbf{T}=\mathbf{F}\mathbf{F}^T$, where $\mathbf{F}=[\mathbf{q}_1\,\mathbf{q}_2\cdots\mathbf{q}_N]$ is a $d\times N$ matrix ($N\leqslant d$). We note that the output of the quadratic filter can be expressed as

$$\varphi=\mathbf{z}^T\mathbf{T}\mathbf{z}=\mathbf{z}^T\mathbf{F}\mathbf{F}^T\mathbf{z}=\mathbf{v}^T\mathbf{v}, \tag{8}$$

where $\mathbf{z}$ is an $r\times c$ subregion of the image and $\mathbf{v}=[v_1\,v_2\cdots v_N]^T$ is a vector of the projection of $\mathbf{z}$ on the N columns of $\mathbf{F}$. Clearly, $\mathbf{v}$ is a function of the spatial region of the image that is represented by $\mathbf{z}$. The question is how

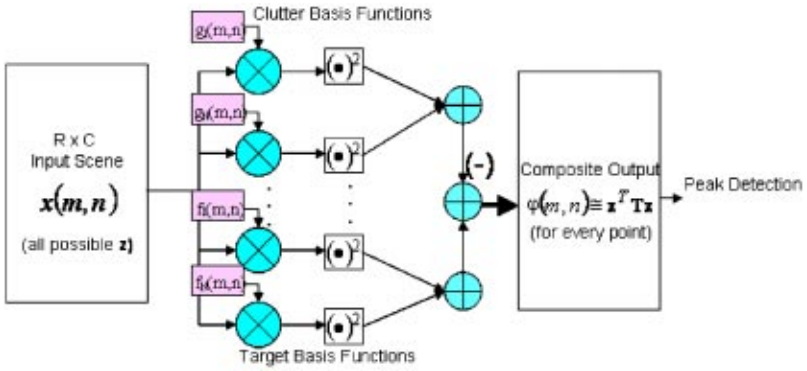

Fig. 1 QCF filter architecture.

the elements of **v** can be computed over the entire scene using efficient signal processing methods. Let $z(m,n)$ represent the full image to be processed. Furthermore, let us reorder the elements of $\mathbf{q}_i$ into a $r\times c$ mask $f_i(m,n)$. It is then easy to see that the values of v_i (i.e., i'th element of **v**) for all locations within the image can be obtained via the 2-D correlation of $f_i(m,n)$ and input test image $z(m,n)$ as

$$v_i(m,n)=z(m,n)\otimes f_i(m,n), \quad 1\leqslant i\leqslant N, \tag{9}$$

where the symbol $\otimes$ indicates the 2-D correlation operation. Here, we have expressed v_i along with the indices (m,n) to indicate that it is a function of the location of **z** in the image. The samples of $v_i(m,n)$ can be viewed as a partial result at each point in the image obtained by projecting the $r\times c$ region of the image centered at that point on $\mathbf{q}_i$. To obtain the quadratic term in Eq. (8) for all points in the image, it now remains to square the pixel surfaces of partial results and add them. Therefore, the output of the quadratic filter for all points in the image is given by

$$\varphi(m,n)=\sum_{i=1}^{N}|v_i(m,n)|^2=\sum_{i=1}^{N}|z(m,n)\otimes f_i(m,n)|^2. \tag{10}$$

The expression in Eq. (10) is a succinct way to express the output of the quadratic filter in response to the full input image. The computations can be readily dealt with using N 2-D cross-correlation operations, each of which is efficiently implemented using fast Fourier transforms (FFTs). In fact, a more general form that takes eigenvectors with negative eigenvalues into account is

$$\varphi(m,n)=\sum_{i=1}^{N_2}|z(m,n)\otimes f_i(m,n)|^2 - \sum_{i=1}^{N_1}|z(m,n)\otimes g_i(m,n)|^2, \tag{11}$$

where $f_i(m,n)$ and $g_i(m,n)$ are eigenvectors corresponding to positive and negative eigenvalues, each reshaped into an $r\times c$ mask.

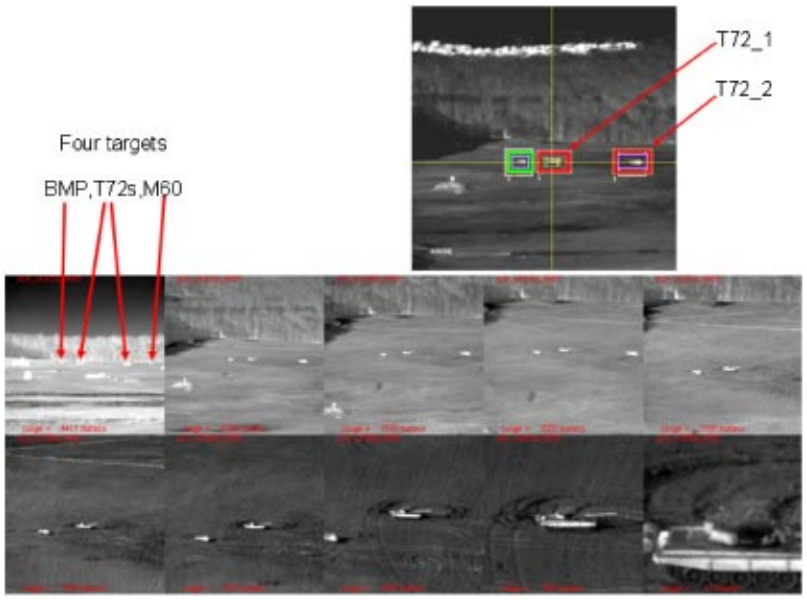

Fig. 2 Image sequence 1.

Figure 1 depicts the architecture required to efficiently process a full image with the quadratic filter. Essentially, the input image is correlated with a bank of linear filters to obtain partial results that are squared and added to obtain the desired quadratic output for every point in the input image. Although several branches of linear filters are required, the architecture shown comprises a single quadratic filter implementation. Since each of the branches can be implemented using FFTs, the number of multiplications required in the architecture is proportional to $N\log(rc)$, where N is the number of branches, and $r\times c$ is the size of the filter masks. This QCF architecture is general, in that any set of basis functions can be implemented, in this case specifically the TBF, not just those derived using the methodology outlined in this work.

4 Detection Metrics

Two detection metrics are used, the first being the peak-to-side lobe ratio (PSR) of the detected objects used for the MACH filter results. The PSR makes use of the standard z-score type normalization on the correlation surface, where each PSR value is based on the local target-sized area statistics, as shown in Eq. (12), where μ is the local mean, σ is the local standard deviation, and τ is the original correlation value at each coordinate.

$$\mathrm{PSR}\equiv(\tau-\mu)/\sigma. \tag{12}$$

The PSR is a measure of the sharpness of the correlation peak that has been shown to be maximized by the MACH filter design process,[12] and is therefore a good metric to assess the performance of the MACH filter. On the other hand, PSR is not optimized by the QCF design algorithm, and is therefore not suitable for measuring QCF performance. For now, we use the composite output [a confidence value (Conf) created by the composite output of the architecture in Fig. 1] as the QCF detection metric. In general, however, it may be desirable to apply a normalization to the QCF output to allow meaningful comparisons across widely varying conditions.

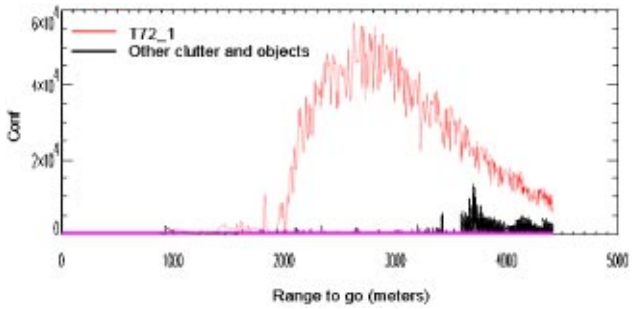

Fig. 3 TBF confidence of T72_1 trained from 4415 to 2695 meters.

The PSR and Conf measures produce results on different scales, but since both show the relative detection values on the desired target and other objects, the salient robustness can be discerned directly for comparison purposes.

5 Initial Dataset

Several samples of the initial dataset are shown in Fig. 2, which also points out the multiple targets in the sequence. Each additional test data sequence is similar but uses different targets in different locations having consequently different backgrounds.

6 Quantifying Performance Results

To illustrate the margin for discrimination, we choose to use strip chart plots showing detection metric amplitude versus range-to-go, which exemplifies the magnitude of the difference between the desired target detectability and other objects, and nondesired targets that are potentially detected. We selected this type of chart, since it depicts robustness in a much more intuitive way than typical receiver operating curves (ROC), which have been noted for misinterpretation.[13]

Figure 3 shows the confidence value plotted on the TBF filter encoding of the primary T72_1 target (the T72 tank on the left in Fig. 2), showing in red confidence values of the

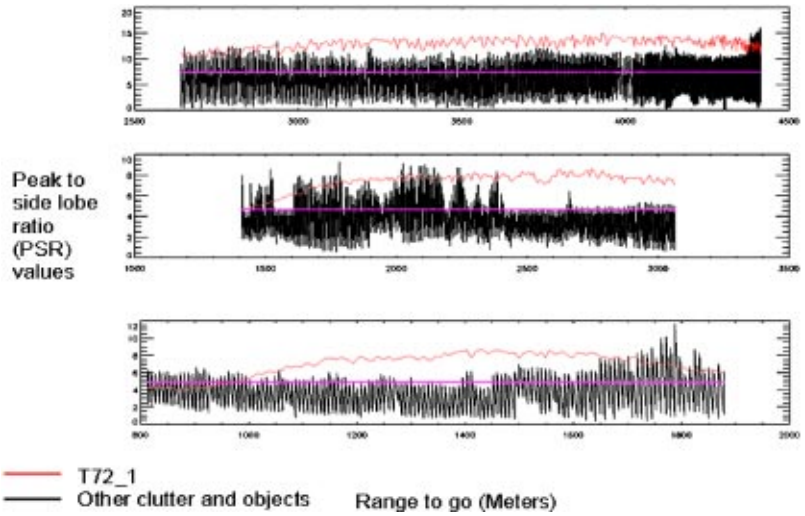

Fig. 4 Three MACH filter results trained on T72_1 at their respective ranges.

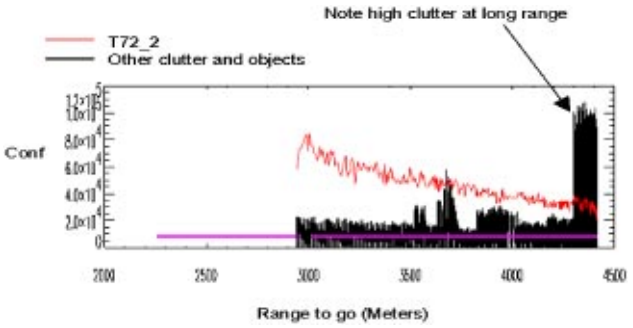

Fig. 5 TBF results on T72_2 target.

target detections and in black the other objects and clutter detected. At all ranges, the target has a higher confidence value, however, between 4400 and 3600 m, there is reduced detection margin than ranges below 3500 m until the filter is out of its useful range. The clutter is suppressed quite dramatically below 3500 m range to go.

Figure 4, in contrast, shows the PSR value of the linear MACH filter results on three different MACH filters encoded with the same primary T72 target. Although the PSR and confidence value scales are different, it is clear the clutter detections are more prevalent and the discrimination margin is generally less, and at some ranges nonexistent.

Figure 5 shows the confidence value plotted on the TBF filter encoding of the secondary T72_2 target (the T72 tank on the right in Fig. 2), again showing in red the target detections and in black the other objects detected. The clutter at long ranges clearly has significantly higher confidence values. This T72 target goes out of the field of view at 2700 m range to go, but note there is little clutter of consequence at the closer ranges.

Figure 6 shows the MACH PSR for comparison to the TBF results in Fig. 5. Again, there is little or no robustness margin in the MACH detection results.

Another dataset with a different target is shown in Fig. 7. Note the TBF results in Fig. 8 have clutter confidence values that exceed the target value on several specific frames. These frames are unique, in that the TBF results are distinctly different than the other frames due to a response to some noise. This noise can potentially be preprocessed out, however, it was left in here to emphasize the sensitivity of the TBF to certain noise.

Figure 9 is the same data that is plotted in Fig. 8 with the vertical scale changed such that the TBF target confidence value response is shown more clearly. Comparing Fig. 9 to Fig. 10 over the same range, there is clutter/noise response, which corresponds for both the TBF and MACH.

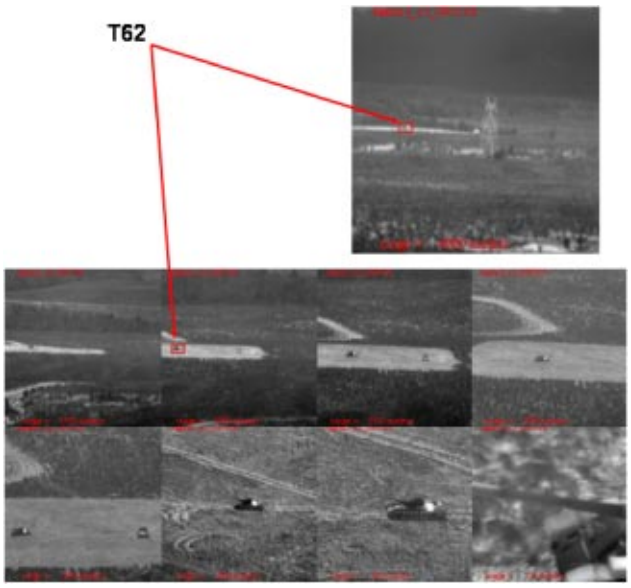

Fig. 7 Image sequence 2.

The margin for discrimination with the TBF is shown to be better than the MACH overall on the T62 target, but the MACH results show a healthy margin also.

7 Summary and Conclusions

We discuss the theory and implementation of QCF filters, and show performance on the tuned basis function with strip chart plots delineating the robustness margin over varying ranges with several different datasets. We also quantify the MACH filter performance over the same data, illustrating the difference in robustness with the TBF.

There are two main advantages of the QCF approach. First, there is only one composite correlation surface to be searched. This is simpler (both conceptually and computationally) than individually assessing multiple correlation surfaces created by a bank of linear correlation filters. The process of selecting the best match among the linear correlation surfaces can also be prone to errors. The second advantage is that a single QCF design can encompass much larger amounts of variations in the target signatures. Although the QCF's computational requirements are comparable to a bank of linear filters, the basis functions form natural partitions that are best suited for optimizing the

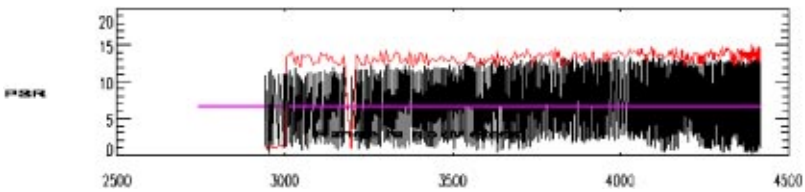

Fig. 6 MACH filter results on T72_2.

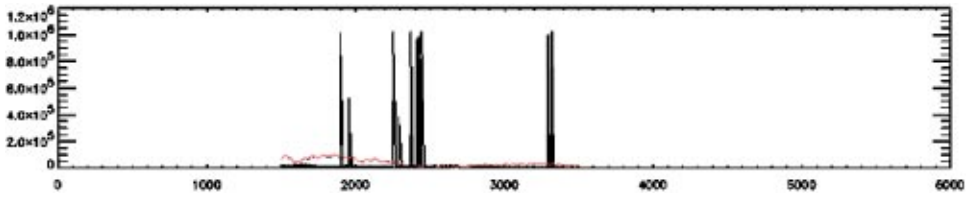

Fig. 8 TBF results on T62 (Image sequence 2).

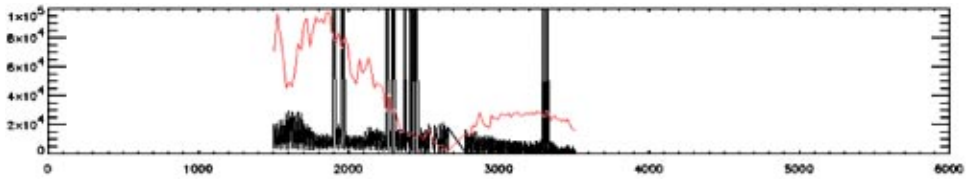

Fig. 9 TBF results (rescaled) on T62 (Image sequence 2).

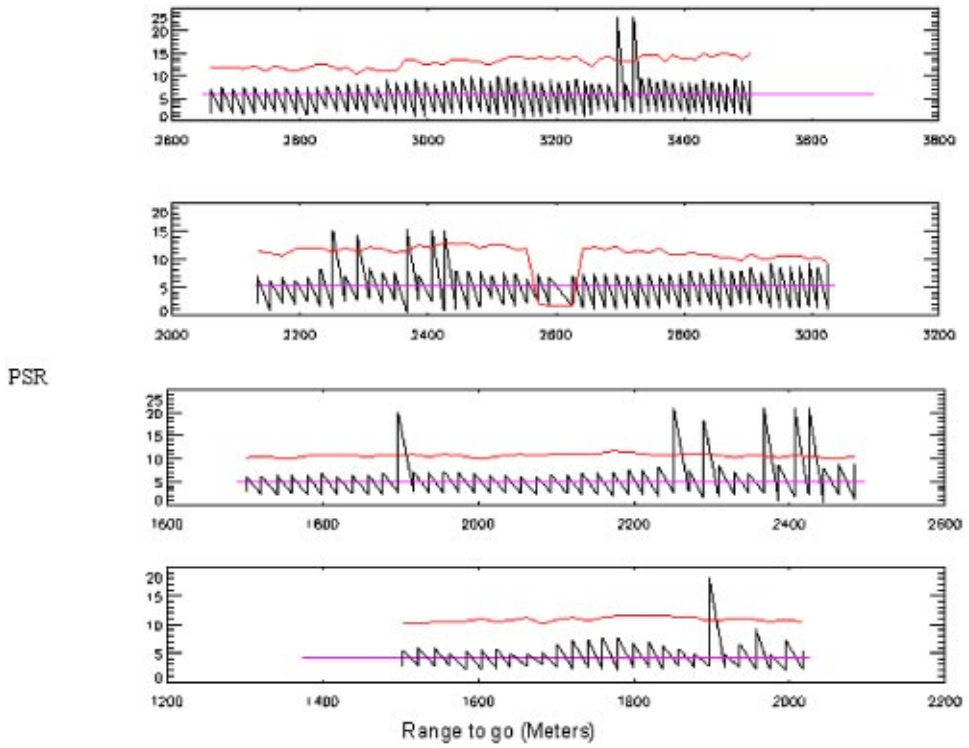

Fig. 10 MACH results on T62.

QCF performance metric (Conf). With the TBF approach, a filter breakdown by range or orientation is not required or is certainly less constraining.

The comparison tests are staged such that the linear filters had every advantage, including smaller range requirements than the TBF filters. This is done to illustrate the robustness margin of the newer TBF filter encoding method. The TBF filters in every case have performed better consistently over the test data used, however, we also show cases where the TBF does not exceed the original linear MACH filters.

8 Future Work

We believe the test data provided some variation in target and background signatures, however, we plan continued tests on more varied data to fully characterize the detection performance for the TBF filters. Larger test sets that represent a broader range of conditions are essential for greater statistical significance, and a better understanding of the merits of the QCF detection approach.

The image pixel detection coordinates of the MACH and TBF are shown to be different during our testing. Subsequent processing of the detected regions of interest using a classification postprocessor will help clarify any additional performance impacts of the TBF filters.

References

1. J. Starch, R. Sharma, and S. Shaw, ''A unified approach to feature extraction for model based ATR,'' *Proc. SPIE* **2757**, 294–305 (1996).
2. D. Casasent and R. Shenoy, ''Feature space trajectory for distorted object classification and pose estimation in SAR,'' *Opt. Eng.* **36**(10), 2719–2728 (1997).
3. B. Bhanu and J. Ahn, ''A system for model-based recognition of articulated objects,'' *Proc. Intl. Conf. Patt. Recog.*, pp. 1812–1815 (1998).
4. S. Z. Der, Q. Zheng, R. Chellappa, B. Redman, and H. Mahmoud, ''View based recognition of military vehicles in LADAR imagery using CAD model matching,'' in *Image Recognition and Classification, Algorithms, Systems and Applications*, B. Javidi, Ed., pp. 151–187, Marcel Dekker, Inc., New York (2002).
5. L. A. Chan, S. Z. Der, and N. M. Nasrabadi, ''Neural based target detectors for multi-band infrared imagery,'' in *Image Recognition and Classification, Algorithms, Systems and Applications*, B. Javidi, Ed., pp. 1–36, Marcel Dekker, Inc., New York (2002).
6. D. Torreiri, ''A linear transform that simplifies and improves neural network classifiers,'' *Proc. Intl. Conf. Neural Net.* **3**, 1738–1743 (1996).
7. H. S. Ranganath, D. E. Kerstetter, and S. R. F. Sims, ''Self partitioning neural networks for target recognition,'' *Neural Networks* **8**(9), 1475–1486 (1995) and *Applications of Neural Networks in Electromagnetics*, C. Christodoulou and M. Georgiopoulos, Eds., pp. 104–108, Artech House, Inc., Boston (2001).
8. B. V. K. Vijaya Kumar, ''Tutorial survey of composite filter designs for optical correlators,'' *Appl. Opt.* **31**, 4773–4801 (1992).
9. A. Mahalanobis, B. V. K. Vijaya Kumar, and S. R. F. Sims, ''Distance classifier correlation filters for multi-class target recognition,'' *Appl. Opt.* **35**, 3127–3133 (1996).
10. A. VanderLugt, ''Signal detection by complex spatial filtering,'' *IEEE Trans. Inf. Theory* **10**, 139–145 (1964).
11. A. Mahalanobis, B. V. K. Vijaya Kumar, S. R. F. Sims, and J. Epperson, ''Unconstrained correlation filters,'' *Appl. Opt.* **33**, 3751–3759 (1994).
12. A. Mahalanobis and B. V. K. Vijaya Kumar, ''On the optimality of the MACH filter for detection of targets in noise,'' *Opt. Eng.* **36**(10), 2642–2648 (1997).
13. S. R. F. Sims, ''Putting ATR performance on an equal basis: The measurement of knowledge base distortion and relevant clutter,'' *Proc. SPIE* **4050**, 55–60 (2000).

S. Richard F. Sims is a research engineer at the U.S. Army Research, Development, and Engineering Command, Aviation and Missile Research, Development and Engineering Center. His target acquisition group implemented the first tactical autonomous target identification demonstrated in missile flight. He has been involved in the conceptualization and development of signal, image, and data processing algorithms in hardware and software for target acquisition and tracking for missile seeker and missile fire control applications for more than two decades. He received his BSE, MSE, and PhD, in electrical and computer engineering at the University of Alabama in Huntsville.

Abhijit Mahalanobis is a principle research engineer at Lockheed Martin, Orlando, and is currently the technical lead for ATR programs in the research and technology group. His main interests are in multisensor automatic target recognition, pattern recognition, and image processing. He has more than 100 journal and conference publications in this area. He has pioneered new approaches in the field of correlation pattern recognition, which are being pursued by many research groups in the U.S. and abroad. He completed his BS degree with honors at the University of California, Santa Barbara, in 1984. He then joined the Carnegie Mellon University and received MS and PhD degrees in 1985 and 1987, respectively. Prior to joining Lockheed Martin, he worked at Raytheon (formerly Hughes) in Tucson, and was a faculty member of the University of Arizona.