Comparative study of maximum average correlation height filter variants using ladar imagery

Alan Van Nevel
Naval Air Warfare Center
China Lake, California 93555
E-mail: vannevelaj@navair.navy.mil

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

Abstract. We present a study of a family of maximum average correlation height (MACH) filters. MACH filters were introduced by Mahalanobis et al., and several modifications such as the extended MACH (EMACH) and generalized MACH (GMACH) have been introduced to enhance the utility of the MACH filter approach. A comparison between the different filtering approaches and processing techniques is presented for the specific case of laser radar (ladar) imagery. The comparison utilizes both synthetic data for training and testing in one case, and synthetic training data and collected ladar imagery for testing in the second case. The results indicate that the GMACH variant of MACH filters is superior to both MACH and EMACH filters for the case of laser radar. [DOI: 10.1117/1.1531975]

Subject terms: ladar; correlation filters; target recognition.

Paper 020029 received Jan. 29, 2002; revised manuscript received Jul. 1, 2002; accepted for publication Jul. 8, 2002.

1 Introduction and Background

Laser radar (or ladar) sensors have been receiving increased attention in recent years. As these sensor systems continue to improve, a small, low cost tactical ladar may become feasible. The desirability of ladar sensors arises from several different factors. First and foremost, the nearly three-dimensional nature of the data and the high resolution of detail that ladar sensors can provide holds great promise for improved capabilities such as target/object recognition, terrain navigation, and aimpoint selection. In addition, while ladar sensors are active systems, the emissions are much more focused and of shorter range, making them more difficult to detect. All of these capabilities suggest that ladar sensors could be a useful tool for detecting and classifying targets and objects in cluttered backgrounds.

Correlation filters, and in particular maximum average correlation height (MACH) filters or optimal trade-off synthetic discriminant filters (OTSDF),[1–8] have been used to detect and classify objects in imagery. Studies using MACH filters to detect and classify military vehicles in cluttered environments[1,9] have demonstrated high detection and classification rates. The MACH filters introduced better distortion tolerance than what could be achieved with earlier correlation filters,[2] which leads to improved performance and reduced computational load (due to the reduced filter count needed to cover all aspect angles).

The class of MACH filters (which includes OTSDFs) is chosen for feature detection for several reasons. As discussed, the filters can incorporate varying degrees of distortion tolerance and can be built to generalize classes of targets. Another benefit of the algorithm is that it is statistically optimum and depends on a realistic, mathematically rigorous optimization procedure as opposed to other heuristic methods. A final consideration is the computational efficiency: the MACH filters require no segmentation or edge detection preprocessing and the correlation step can be performed rapidly using dedicated FFT hardware. In this work, the term MACH filters is used as a general term to describe all three variants, MACH, GMACH, and EMACH.

Unfortunately, the response of MACH filters to nontraining (clutter or confuser objects) images can be unacceptably high, resulting in high false alarm rates. Recent work by Alkanhal et al.[10,11] introduced the extended MACH (EMACH) filter, which was designed to reduce the reliance on the mean training image in the MACH filters and improve clutter rejection. Another improvement over MACH filters, the generalized MACH (GMACH), was suggested by Mahalanobis et al.[2] to minimize the variance in correlation peaks among the training images. Those images in the training set that closely resembled the mean image have significantly larger correlation peaks than images at the ''edges'' of the training set. The GMACH filter attempted to address this problem.

In this work, a comprehensive comparison of MACH, GMACH, and EMACH filters is presented using synthetic imagery to train the filters, while both synthetic and collected ladar imagery is used to test the filters. The primary metric used for comparison is the Fisher ratio.[12]

2 MACH, GMACH, and EMACH Filters

2.1 *Introduction and Theory*

Mahalanobis, Kumar, et al.[1–8] have recently introduced an advanced family of correlation filters known as maximum average correlation height (MACH) filters. The use of MACH filters for automatic target recognition (ATR) has been motivated by the properties of the MACH filters.

Since MACH filters are a correlation filter, they can be implemented easily and operate at high rates of speed given current hardware processing speeds.

In the discussion of the MACH (including GMACH and EMACH) filters that follows, bold lowercase indicates a column vector, while bold uppercase represents a diagonal matrix. The strict MACH filter results from maximizing the ratio

$$\mathrm{J}(\mathbf{h})=\frac{|\mathbf{h}^{+}\mathbf{m}|}{\mathbf{h}^{+}\mathbf{S}\mathbf{h}}, \tag{1}$$

where $\mathbf{h}$ is the correlation filter and $\mathbf{m}$ is the average of the training images that have undergone a 2-D FFT. The symbol + is used to indicate conjugate transpose. Each transformed image is lexicographically ordered to form a vector. $\mathbf{S}$ is the average similarity measure (ASM) matrix

$$\mathbf{S}=\frac{1}{N}\sum_{k=1}^{N}(\mathbf{X}_k-\mathbf{M})(\mathbf{X}_k-\mathbf{M})^{+}. \tag{2}$$

In Eq. (2), $\mathbf{X}_k$ are the individual training images, again in the Fourier domain. The training image is lexicographically ordered and its elements placed on the diagonal of $\mathbf{X}_k$, while $\mathbf{M}$ is the mean training image, arranged similarly to $\mathbf{X}_k$. Furthermore, all of the processing to generate the filters is performed in the Fourier domain to gain translational invariance. It is possible to perform the processing in other domains (e.g., wavelet or spatial), but care must be taken to properly register the training imagery.

The optimal filter $\mathbf{h}$ is then given by

$$\mathbf{h}=\mathbf{S}^{-1}\mathbf{m}. \tag{3}$$

Variants on the MACH filter can be achieved by varying the performance metric one wishes to maximize. For example Refregier[5] has developed optimal trade-off synthetic discriminant filters (OTSDFs), which attempt to minimize the energy functional

$$E(\mathbf{h})=\mathbf{h}^{+}\mathbf{Q}\mathbf{h}-\delta|\mathbf{h}^{+}\mathbf{m}|, \tag{4}$$

where

$$\mathbf{Q}=\alpha\mathbf{P}+\beta\mathbf{D}+\gamma\mathbf{S}. \tag{5}$$

$\mathbf{S}$ is as defined before, $\mathbf{P}$ is the power spectral density of the expected noise [output noise variance (ONV) term], and $\mathbf{D}$ is the average power spectral density [average correlation energy (ACE) term] of the training set. The power spectral density $\mathbf{P}$ and the ACE term $\mathbf{D}$ can be expressed as

$$\mathbf{P}=\frac{1}{M}\sum_{i=1}^{M}\mathbf{Z}_i\mathbf{Z}_i^{+}, \tag{6}$$

$$\mathbf{D}=\frac{1}{N}\sum_{i=1}^{N}\mathbf{X}_i\mathbf{X}_i^{+}, \tag{7}$$

where $\mathbf{Z}_i$ is a sample image (in the Fourier domain) of the expected noise or clutter, and $\mathbf{X}_i$ is defined as in Eq. (2).

The constants α, β, γ, and δ are nonnegative and must satisfy $\alpha^2+\beta^2+\gamma^2+\delta^2=k$, where k is any positive constant. Minimizing $E(\mathbf{h})$ results in

$$\mathbf{h}=\frac{\delta}{2}\mathbf{Q}^{-1}\mathbf{m}. \tag{8}$$

One can achieve the same result within a scaling factor by replacing $\mathbf{S}$ in Eq. (1) with $\mathbf{Q}$

$$\mathrm{J}(\mathbf{h})=\frac{|\mathbf{h}^{+}\mathbf{m}|^2}{\mathbf{h}^{+}\mathbf{Q}\mathbf{h}}. \tag{9}$$

By varying the parameters, one can optimize filter performance for the situation under study. If one sets $\alpha=\beta=0$, the result is the strict MACH filter discussed earlier. Further variations can be made to the basic idea, including the extension to multiple class discrimination using distance classifier correlation filters[3] (DCCFs), which are able to distinguish between multiple classes of similar objects (e.g., T72s versus M1A1 tanks).

To calculate the GMACH filter, one follows a procedure similar to that outlined before. The only change is in the $\mathbf{Q}$ term in the denominator. Assume that there are N training images $\mathbf{x}_i$. The covariance of the correlation peak values is given by

$$\sigma_{\mathrm{peak}}^2=\frac{1}{N}\mathbf{h}^{+}\left[\sum_{i=1}^{N}(\mathbf{x}_i-\mathbf{m})(\mathbf{x}_i-\mathbf{m})^{+}\right]\mathbf{h}=\mathbf{h}^{+}\mathbf{E}\mathbf{E}^{+}\mathbf{h}=\mathbf{h}^{+}\Sigma\mathbf{h}. \tag{10}$$

Here, the columns of $\mathbf{E}$ are the training images minus the average $\mathbf{m}$, and Σ is an estimate of the covariance matrix of the training images (a large matrix of size d^2xN and rank N).

The GMACH filter[2] reduces the variation in the correlation peak values by maximizing the more generalized MACH criterion:

$$J(\mathbf{h})=\frac{|\mathbf{h}^{+}\mathbf{m}|^2}{\mathbf{h}^{+}[\delta\cdot\Sigma+\mathbf{Q}]\mathbf{h}}, \tag{11}$$

where $\mathbf{Q}=\alpha\mathbf{P}+\beta\mathbf{D}+\gamma\mathbf{S}$ is a diagonal matrix combining the ACE, ASM, and ONV metrics, and δ is a scalar used to control how much the peak value variations should be minimized (i.e., another optimal trade parameter). The solution is

$$\mathbf{h}_g=[\delta\Sigma+\mathbf{Q}]^{-1}\mathbf{m}. \tag{12}$$

The issue here is how to best compute the inverse of the very large matrix. Remember, $\mathbf{Q}$ is diagonal and Σ is of rank N. Using a technique for matrix inversion from Kailath's book on linear systems,[13] the solution can be shown to be

$$\begin{aligned}\mathbf{h}_g&=[\mathbf{Q}^{-1}-\delta\mathbf{Q}^{-1}\mathbf{E}(\mathbf{E}^{+}\mathbf{Q}^{-1}\mathbf{E}+\mathbf{I})^{-1}\mathbf{E}^{+}\mathbf{Q}^{-1}]\mathbf{m}\\&=\mathbf{Q}^{-1}\mathbf{m}-\delta\mathbf{Q}^{-1}\mathbf{E}(\mathbf{E}^{+}\mathbf{Q}^{-1}\mathbf{E}+\mathbf{I})^{-1}\mathbf{E}^{+}\mathbf{Q}^{-1}\mathbf{m}.\end{aligned} \tag{13}$$

Fortunately, in this form all matrices to be inverted are only of size N and the inversion of the large matrix can be entirely avoided. This version of the MACH filter yields much better distribution of peak values (and indirectly PSR values) across all training images. For larger values of δ, the peak values become exactly the same for all training images. The conventional MACH filter is obtained as δ goes to zero. The EMACH filter, as introduced by Alkanhal, Kumar, and Mahalanobis,[10] is designed to reduce the emphasis on the mean training image while preserving the distortion tolerance of the original MACH filters. To do so, the ASM metric in Eq. (2) is modified, and a new metric, the all image correlation height (AICH) is introduced. The AICH metric is given by

$$\mathbf{C}^{\delta}=\frac{1}{N}\sum_{i=1}^{N}(\mathbf{x}_i-\delta\mathbf{m})(\mathbf{x}_i-\delta\mathbf{m})^{+}, \tag{14}$$

where $0\leqslant\beta\leqslant 2$. The AICH reduces the reliance on the mean training image and reduces the contribution of low frequency components from the average image. Using the same reasoning as in the MACH implementation,[2] we wish to minimize the mean squared error of the training correlation planes compared to an ideal correlation plane. The ideal correlation plane is given by[10]

$$\mathbf{f}_{\mathrm{opt}}=(1-\delta)\mathbf{M}\mathbf{h}^{*}. \tag{15}$$

The ASM measure in Eq. (2) is then modified to measure the dissimilarity of the training images to $(1-\delta)\mathbf{M}\mathbf{h}^{*}$

$$\mathbf{S}^{\delta}=\frac{1}{N}\sum_{i=1}^{N}[\mathbf{X}_i-(1-\delta)\mathbf{M}]^{+}[\mathbf{X}_i-(1-\delta)\mathbf{M}]. \tag{16}$$

This new measure provides better clutter rejection than the original MACH filter by emphasizing the higher frequency components. Maximizing the AICH and minimizing the modified ASM leads to the EMACH criterion

$$\mathrm{J}^{\delta}(\mathbf{h})=\frac{\mathbf{h}^{+}\mathbf{C}^{\delta}\mathbf{h}}{\mathbf{h}^{+}(\mathbf{P}+\mathbf{S}^{\delta})\mathbf{h}}. \tag{17}$$

The $\mathbf{P}$ term is the ONV term, which presumes an additive noise and helps to maintain noise tolerance. $\mathbf{P}$ reduces to the identity if one considers white noise with unit variance. The new EMACH filter is found from

$$(\mathbf{P}+\mathbf{S}^{\delta})^{-1}\mathbf{C}^{\delta}\mathbf{h}=\lambda\mathbf{h}, \tag{18}$$

where λ is a scalar equal to $\mathrm{J}^{\delta}(\mathbf{h})$. Therefore, $\mathbf{h}$ must be an eigenvector of $(\mathbf{P}+\mathbf{S}^{\delta})^{-1}\mathbf{C}^{\delta}$ with eigenvalue λ. Typically, the largest eigenvalue and corresponding eigenvector is used as the EMACH filter. It is unknown if smaller eigenvalues will provide better discriminatory performance. Computing the eigenvectors is identical to the method used in computing distance classifier correlation filters.[3] (Another variant, the eigen-EMACH has been proposed,[11] which is not considered here.) A simple modification to Eq. (17) can be made, replacing the term in parentheses, with a modified $\mathbf{Q}$

Fig. 1 Synthetic range image of M60 tank model. Range to target is approximately 850 m.

$$\mathbf{Q}^{-1}\mathbf{C}^{\delta}\mathbf{h}=\lambda\mathbf{h}. \tag{19}$$

This modification adds in the freedom to weight the three terms in a similar fashion as the MACH and GMACH filters.

3 Training and Testing Approach

3.1 *Training Data*

This study uses synthetic ladar imagery as well as collected real world data. High quality imagery can be easily produced using high fidelity CAD/CAM models of targets of interest. An example image can be seen in Fig. 1. To generate a synthetic ladar range image, the seeker-relative target aspect and range needs to be specified, as well as the resolution of the image. The relative aspect can be defined by (r,θ,φ) with the coordinate origin at the center of the target. This choice reduces the number of degrees of freedom in the choice of target aspect by assuming the target is resting on locally flat and horizontal ground. The flat ground assumption eliminates any variability in the relative pitch of the target and would not be valid for a tactically fielded ATR system, but the choice suffices for the present study.

A common problem encountered in processing range imagery is the need to segment the image, separating possible targets from the background. The quality of the segmentation can adversely impact the accuracy of ATR algorithms. In an attempt to improve correlator performance without incurring the computational cost associated with segmentation algorithms, a rendering technique[14,15] is applied to the range imagery as a preprocessing step. A range image can be thought of as a surface, with each pixel having an estimate of the local surface normal $\mathbf{S}(i,j)$, where (i,j) indicates pixel locations. A new rendered intensity image can be formed by taking the inner product of $\mathbf{S}(i,j)$ with an illumination function $\mathbf{L}$.

$$\mathbf{L}=l_x\mathbf{i}+l_y\mathbf{j}+l_z\mathbf{k}, \tag{20}$$

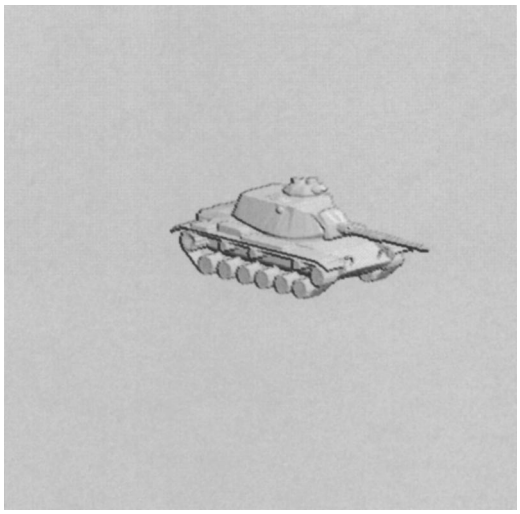

Fig. 2 Rendered image of synthetic image of Fig. 1, using a virtual light source.

$$\mathbf{S}=s_x\mathbf{i}+s_y\mathbf{j}+s_z\mathbf{k}, \tag{21}$$

$$\mathbf{R}=\mathbf{L}\cdot\mathbf{S}. \tag{22}$$

The result **R** can be interpreted as the intensity image that would result given a light source and a uniform reflectance for all pixels in the image. In addition, the dot product also lends insight to how the surfaces are oriented with respect to the virtual light source.

An example of the rendered image can be seen in Fig. 2. All range data used in this study are preprocessed in this manner. It has been shown[16] that this preprocessing improves correlation-based detection performance. One possible explanation is depicted in Fig. 3. If a target is positioned on flat ground, the range image will display a near constant gradient as one progresses from the bottom of the range image to the top. Figure 3(a) shows a plot of one column of the range data, through the target, while Fig. 3(b) shows a plot of one column that does not include the target. The bump or discontinuity in the near linear range profile is the target location with a small discontinuity at the front of the target, and a large discontinuity at the rear of the target. In Fig. 3(c), the FFT of the 1-D range profile

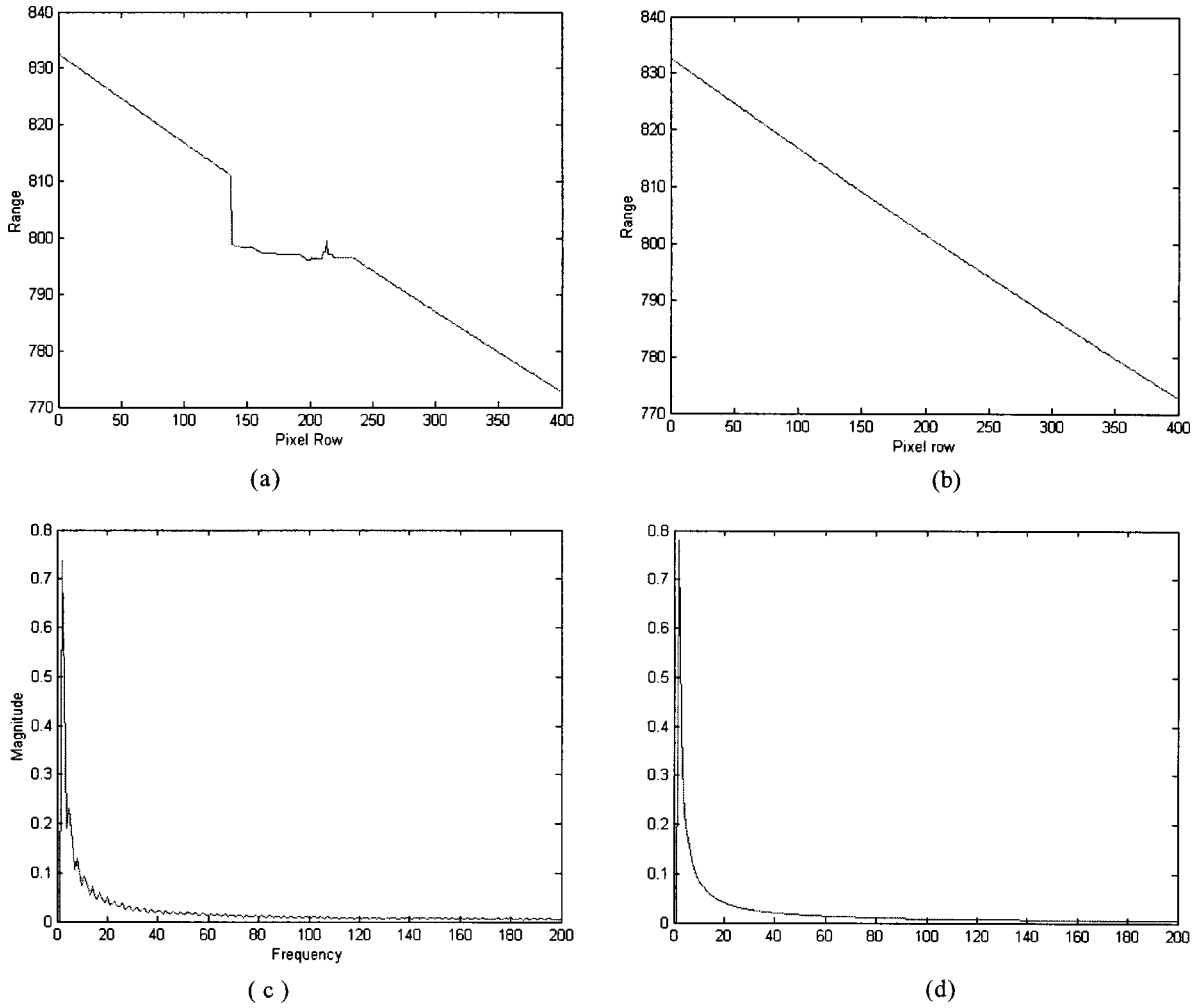

Fig. 3 In (a) and (b), a plot of range values for a single row of the M60 tank in Fig. 1 are shown. The target is present in (a) and absent in (b). In (c) and (d) the magnitude of the FFT of (a) and (b) are shown.

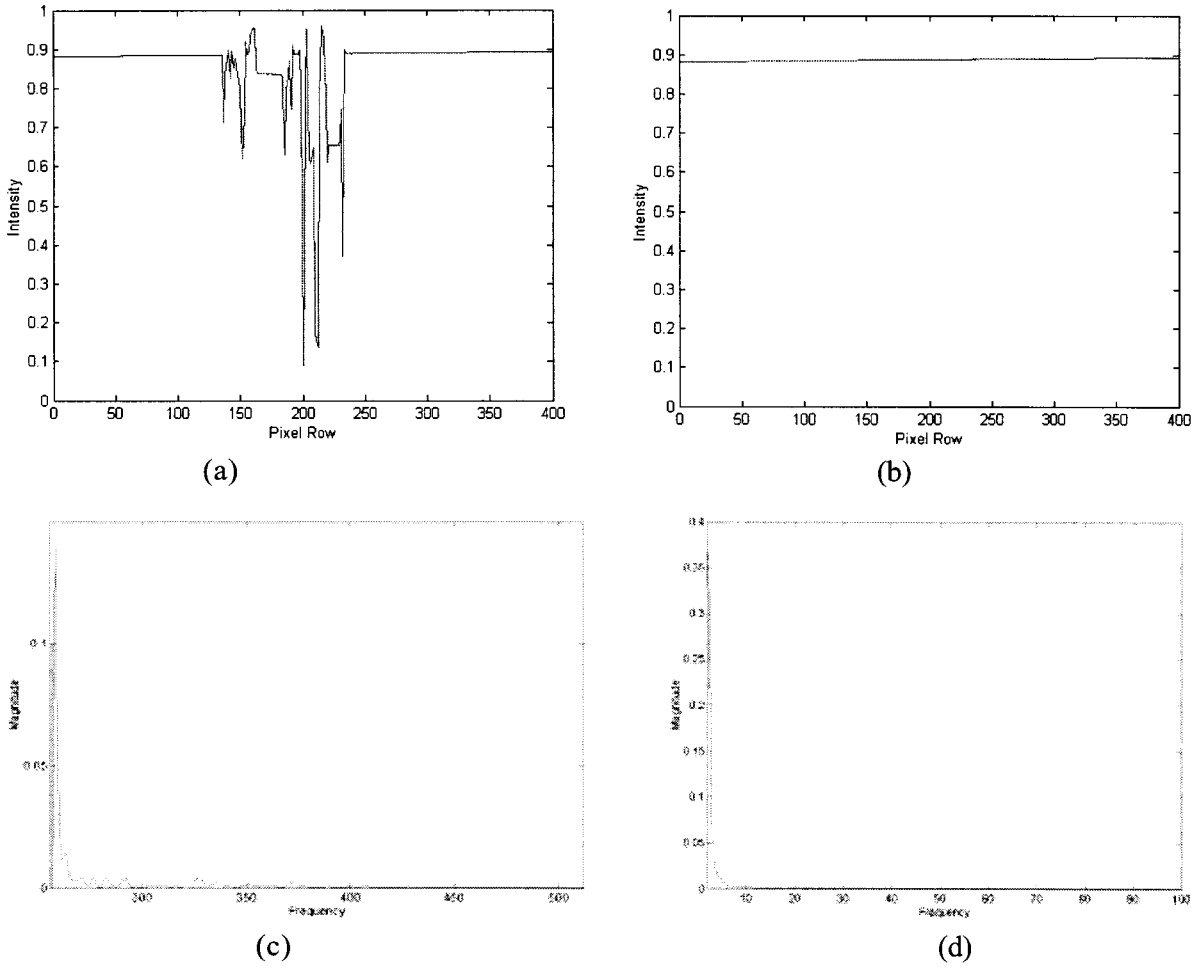

Fig. 4 Plots of the intensity values of the rendered synthetic images for a single row of pixels. (a) shows a row containing the target, while (b) does not. In (c) and (d) the FFT of (a) and (b) are shown.

through the target is shown, while Fig. 3(d) shows the FFT of the data shown in Fig. 3(b). The small discontinuity displays itself as a small perturbation of the linear profile, suggesting that the background needs to be removed in some fashion. The rendering process transforms the linear background to a constant, but this does not significantly impact Fourier-based processing (such as correlation filtering). Plots of the preprocessed data corresponding to the data in Figs. 3(a) to 3(b) can be seen in Figs. 4(a) and 4(b). Note that the background areas are nearly constant.

3.2 *Metrics*

The principal metric used for comparison in this study is the Fisher ratio.[12] The ratio is indicative of the ability of correlation filters to discriminate between the desired class and clutter. The Fisher ration (FR) is given by

$$\mathrm{FR}=\frac{(m_1-m_2)^2}{\sigma_1^2+\sigma_2^2}, \tag{23}$$

where m_1 and σ_1 are the peak to sidelobe ratio (PSR) mean and variance of class one, while m_2 and σ_2 are the mean and variance for class two. The peak to sidelobe ratio is given by

$$\mathrm{PSR}=\frac{\text{peak-mean}}{\text{standard deviation}}=\frac{p-\mu}{\sigma}, \tag{24}$$

where p is the peak of the correlation output, and μ and σ, respectively, are the mean and standard deviation of the correlation output, calculated either over the entire correlation plane, or within a selected window. The benefit of this approach is that the ratio is independent of illumination and amplification effects. The overall peak height can be affected by constant amplification but the ratio will remove this problem. This metric works well in rejecting false peaks due to clutter, since most correlation surfaces for clutter images will not contain a high percentage of energy in a localized window. The Fisher ratio captures key elements of a good correlation filter: Large (PSR) separation between classes, with small (PSR) variations within classes.

3.3 *Synthetic Data Only*

Two different approaches to training and testing the various types of filters are considered. In one case, synthetic imagery is created to train and test the filters, while the second comparison involves synthetic data for training and collected ladar imagery for testing. Care is taken in the

Table 1 Examples of parameters used to create synthetic training and test data. MATLAB notation has been used, e.g., 950:50:1050 is equivalent to a set {950,1000,1050}. Ranges are in meters, aspect is in degrees, with 0 deg corresponding to a head-on view; elevation is in degrees, with 0 deg being horizontal; and a negative elevation indicates the sensor position is above the target.

		Training	Testing
	Range	950:50:1050	925:50:1075
A1	Aspect	75:3:105	70:4:110
	Elevation	−12:−3:−18	{−10,−13.5,−16.5,−20}
	Range	950:50:1050	925:50:1075
A2	Aspect	30:6:60	25:4:65
	Elevation	−12:−3:−18	{−10,−13.5,−16.5,−20}
	Range	800:25:850	787:25:862
A3	Aspect	−15:3:15	−20:4:20
	Elevation	−12:−3:−18	{−10,−13.5,−16.5,−20}
	Range	950:50:1050	925:50:1075
A4	Aspect	−15:3:15	−20:4:20
	Elevation	−12:−3:−18	{−10,−13.5,−16.5,−20}
	Range	1200:50:1300	1175:50:1325
A5	Aspect	30:6:60	25:4:65
	Elevation	{−2,−4,−6}	{−1,−3,−5,−7}

synthetic-data-only case to present test images that are not in the training set, and also to present test images that are outside of the span of the training data. For example, if the training data aspect angle varied from 75 (near broadside view, front right quarter of target) to 105 deg (near broadside, from the right rear quarter of target) in steps of 3 deg, then the test data would vary from 70 to 110 deg, in steps of 4 deg. This ensures that there will be data outside the training range to test distortion tolerance, and that most test data inside the span will not duplicate training data. The same approach is taken with the elevation angle of the sensor and the range to the sensor, which guarantees no test data duplicated any training data.

Two vehicle models were chosen for comparison, an M60 tank, and an SA-8 surface to air missile launcher. Rather than choosing two tank-like vehicles, two dissimilar vehicles were chosen, as the MACH filters would not have sufficient discrimination ability. Previous work,[1,3,9] as well as the prior experience of the authors, has shown that MACH filters do not provide sufficient discrimination ability. The DCCF variant of MACH filters would be able to discriminate between similar vehicles, e.g., M60 and T72 tanks, but that is not the focus of this work. The focus is to determine which variant, if any, of the MACH type filters provides a better response to true class images, while minimizing the response to vehicles and objects not in the training set.

The next step is to create training and test data for each of the vehicle types using a wide variety of aspects, elevation, and ranges to build the dataset. A few samples of

Table 2 Values used in training filters.

Set #	α	β	γ	δ	μ
1	0.00	1.00	0.00	0.10	0.10
2	0.00	1.00	0.00	0.20	0.10
3	0.00	1.00	0.00	0.30	0.10
4	0.00	1.00	0.00	0.40	0.10
5	0.00	1.00	0.00	0.50	0.10
6	0.00	1.00	0.00	0.60	0.10
7	0.00	1.00	0.00	0.70	0.10
8	0.00	1.00	0.00	0.80	0.10
9	0.00	1.00	0.00	0.90	0.10
10	0.00	1.00	0.00	1.00	0.10
11	0.00	0.20	0.80	0.10	0.10
12	0.00	0.20	0.80	0.20	0.10
13	0.00	0.20	0.80	0.30	0.10
14	0.00	0.20	0.80	0.40	0.10
15	0.00	0.20	0.80	0.50	0.10
16	0.00	0.20	0.80	0.60	0.10
17	0.00	0.20	0.80	0.70	0.10
18	0.00	0.20	0.80	0.80	0.10
19	0.00	0.20	0.80	0.90	0.10
20	0.00	0.20	0.80	1.00	0.10
21	0.00	0.00	1.00	0.10	0.10
22	0.00	0.00	1.00	0.20	0.10
23	0.00	0.00	1.00	0.30	0.10
24	0.00	0.00	1.00	0.40	0.10
25	0.00	0.00	1.00	0.50	0.10
26	0.00	0.00	1.00	0.60	0.10
27	0.00	0.00	1.00	0.70	0.10
28	0.00	0.00	1.00	0.80	0.10
29	0.00	0.00	1.00	0.90	0.10
30	0.00	0.00	1.00	1.00	0.10
31	0.00	0.10	0.10	0.10	0.10
32	0.00	0.10	0.10	0.20	0.10
33	0.00	0.10	0.10	0.30	0.10
34	0.00	0.10	0.10	0.40	0.10
35	0.00	0.10	0.10	0.50	0.10
36	0.00	0.10	0.10	0.60	0.10
37	0.00	0.10	0.10	0.70	0.10
38	0.00	0.10	0.10	0.80	0.10
39	0.00	0.10	0.10	0.90	0.10
40	0.00	0.10	0.10	1.00	0.10
41	1.00	0.10	0.10	0.10	0.00
42	1.00	0.10	0.10	0.20	0.00
43	1.00	0.10	0.10	0.30	0.00
44	1.00	0.10	0.10	0.40	0.00
45	1.00	0.10	0.10	0.50	0.00
46	1.00	0.10	0.10	0.60	0.00
47	1.00	0.10	0.10	0.70	0.00
48	1.00	0.10	0.10	0.80	0.00
49	1.00	0.10	0.10	0.90	0.00
50	1.00	0.10	0.10	1.00	0.00

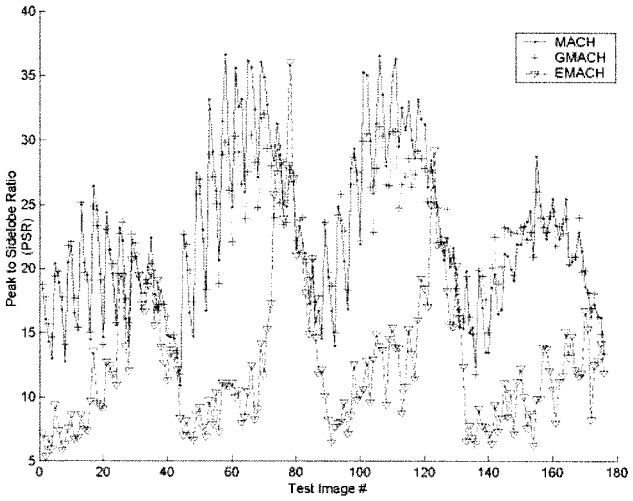

Fig. 5 Plot of PSR versus test images of an M60 tank generated using parameter set A2 from Table 1 for one particular filter parameter set (line 17 Table 2).

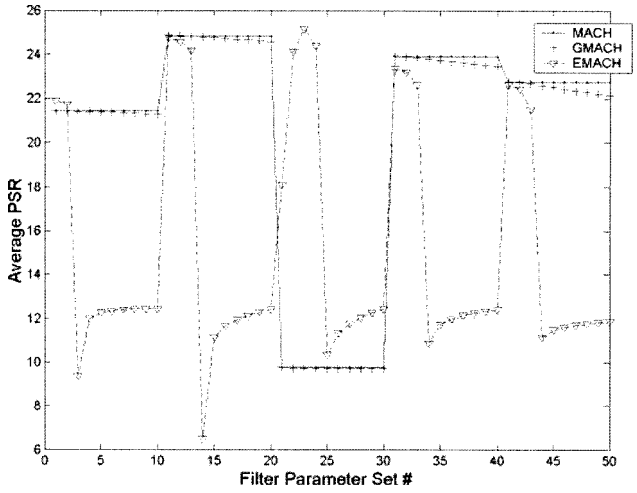

Fig. 7 Average PSR versus filter parameter set number. The average PSR is calculated over 176 test images for a particular filter parameter set. Filters were trained and tested with M60 imagery generated using parameter set A5 from Table 1.

training and test set parameters are displayed in Table 1. As discussed in Sec. 2.1, each of the filter types (MACH, GMACH, and EMACH) have associated parameters that effect their performance. No optimal method for choosing the parameters has been developed to date, and this remains a major drawback to using MACH-type correlation filters. For this comparative study, 50 different parameter sets are considered. These parameters are developed by using five sets of parameters for α, β, and γ, and then varying the δ parameter (used in the GMACH and EMACH filters) from 0.1 to 1 in 0.1 increments, resulting in 50 different parameter sets for GMACH and EMACH filters, but only five for the basic MACH filter, which does not depend on δ. Table 2 shows the filter parameter set. The values for the five parameter sets α, β, and γ have been selected to reflect the different weighting schemes one might consider. For example, $\alpha=0$, $\beta=1$, and $\gamma=0$ corresponds to a MACE filter[2] while $\alpha=0$, $\beta=0$, and $\gamma=1$ is a strict MACH filter. Given a particular training set and the 50 filter parameter sets, 5 MACH filters, 50 GMACH filters and 50 EMACH filters are created.

Following filter generation, the M60 and SA-8 test data are filtered using the MACH, GMACH, and EMACH filters. In Fig. 5 a plot of the PSR for each filter type versus test image number is shown for one filter parameter set (row 17 in Table 2). Figure 5 illustrates the behavior of each of the three filter types using a particular parameter set, and typifies the response over a large set of test images. An alternative way to examine the data is to average the PSR over all test images for a particular filter (needed for the calculation of the FR), and then plot the average PSR versus the filter parameter set number (Figs. 6–8). For example, for x coordinate=17 in Fig. 6, the mean PSR is the mean of all PSR values shown in Fig. 5. Each x coordinate

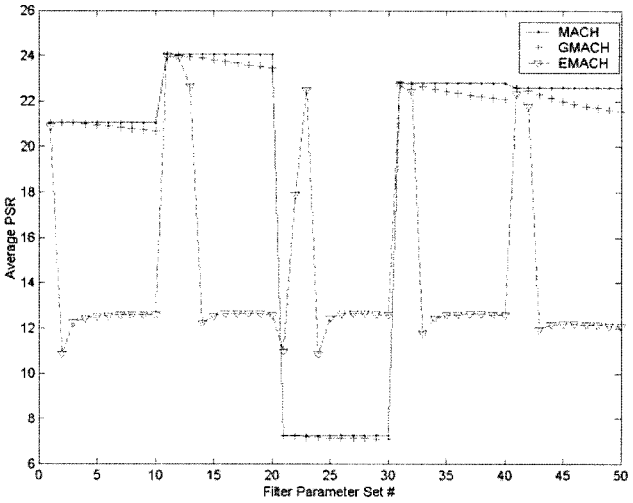

Fig. 6 Average PSR versus filter parameter set number. The average PSR is calculated over 176 test images for a particular filter parameter set. Filters were trained and tested with M60 imagery generated using parameter set A2 from Table 1.

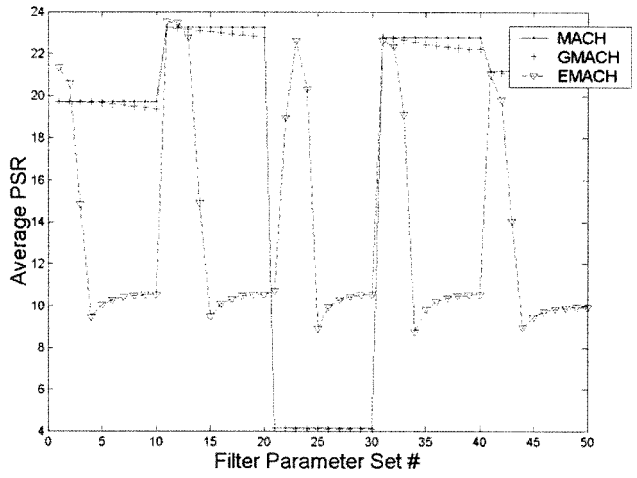

Fig. 8 Average PSR versus filter parameter set number. The average PSR is calculated over 176 test images for a particular filter parameter set. Filters were trained and tested with SA-8 imagery generated using parameter set A2 from Table 1.

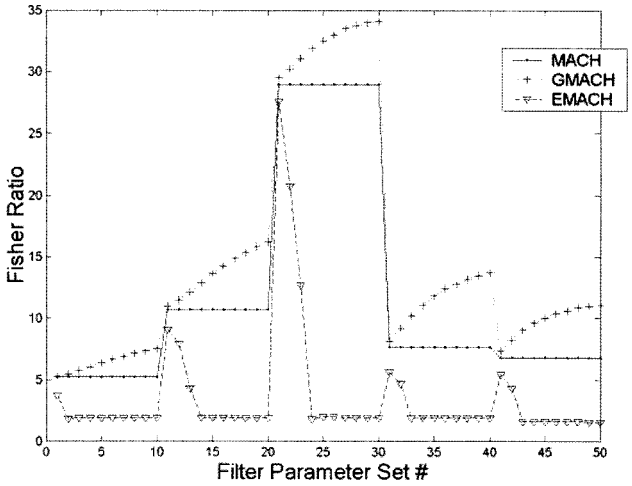

Fig. 9 Fisher ratio versus filter parameter set number. The Fisher ratio was calculated using M60s as true class (generated using set A5), with SA-8s as the false class (generated using set A2), with MACH filters trained using M60 imagery. The different range to target in the two classes was used to ensure the same number of pixels on target.

in Fig. 6 corresponds to a row of parameter values in Table 2, while the y coordinate value is the average PSR of all test images. Figures 6–8 display mean results for selected comparisons using different training and test imagery sets. The Fisher ratio was also calculated for this data. Figures 9–11 show the Fisher ratios versus the filter parameter set number, for a variety of training and test imagery sets. In Figs. 6–11, a flat response over certain regions for the MACH filter is seen. The cause of the flat MACH filter response over certain regions is that no filter parameters changed for that range. For a given α, β, and γ, varying δ will have no impact on PSR values for the MACH filter. The $\{\alpha,\beta,\gamma\}$ set, which displays the highest PSR average, is $\{0,0.2,0.8\}$, while the set that has the lowest PSR performance $\{0,0,1\}$ has the highest Fisher ratio for class separability. One possible explanation for this result is that by

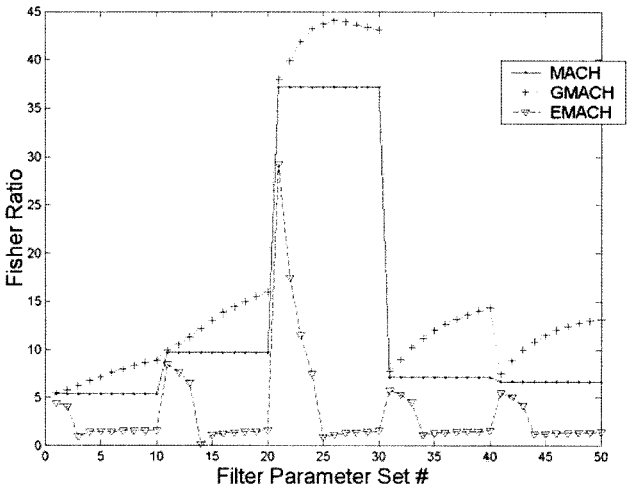

Fig. 10 Fisher ratio versus filter parameter set number. The Fisher ratio was calculated using M60s as true class (generated using set A4), with SA-8s as the false class (generated using set A3), with MACH filters trained using M60 imagery.

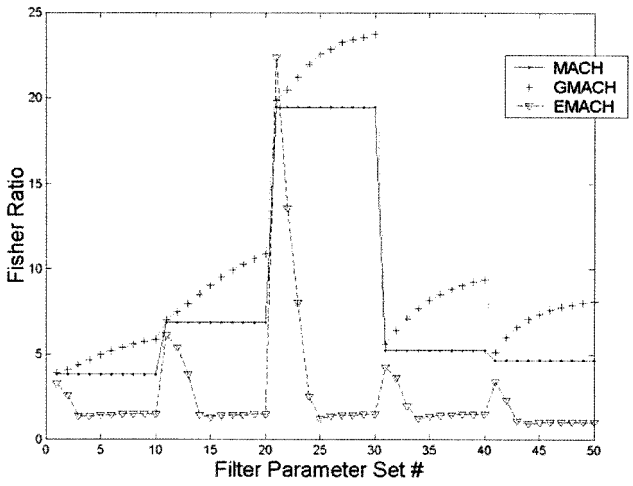

Fig. 11 Fisher ratio versus filter parameter set number. The Fisher ratio was calculated using SA-8s as the true class (generated using set A1), with M60s as the false class (generated using set A1), with MACH filters trained using SA-8 imagery.

using only the ASM term, the method should also lead to smaller in-class PSR variances leading directly to a higher Fisher ratio.

3.4 *Testing with Real Data*

In addition to testing with synthetic data, comparisons between the filter types have been made using the results from tests on data collected from captive flight tests. The dataset used for testing consists of over 440 passes of captive flight tests. Each pass consists of, on the average, ten images. The imagery contains small mobile targets or large fixed targets. The number of targets may vary from as few as four at close ranges to more than 20 at long ranges, and the type of clutter varies from no or light clutter to heavy clutter. The images in Fig. 12 show the types of targets and clutter that are present in the imagery. Since this research effort is geared to algorithm development rather than system evaluation, the dataset used for testing and evaluation is limited to five different passes of the same target site with different approach angles. Within this limited set there are more than 1000 target detection opportunities.

Each of the test images has associated metadata such as estimated aspect angle to target, approximate depression (elevation) angle, and approximate ground truth pixel locations of target vehicles. A wide variety of vehicles can be seen in the imagery, such as M60 tanks, M113 armored personnel carriers, trucks, ZSU-23 antiaircraft artillery, and self-propelled artillery.

To build the training set (and consequently the MACH filters,) some *a priori* information about the test imagery was presumed. As we are not trying to implement MACH-type filters as a detector acting on the entire image, the ground truth data is used to extract image chips, 64×64 pixels, that contain the target vehicles. Other natural clutter image chips are extracted from the same imagery using results from other ladar target detection algorithms.[17] Examples of confusers (nontarget vehicles) and natural clutter chips can be seen in Fig. 13. From the test image (from which we extract the image chip) it is possible to estimate the range to the center of the image and the depression

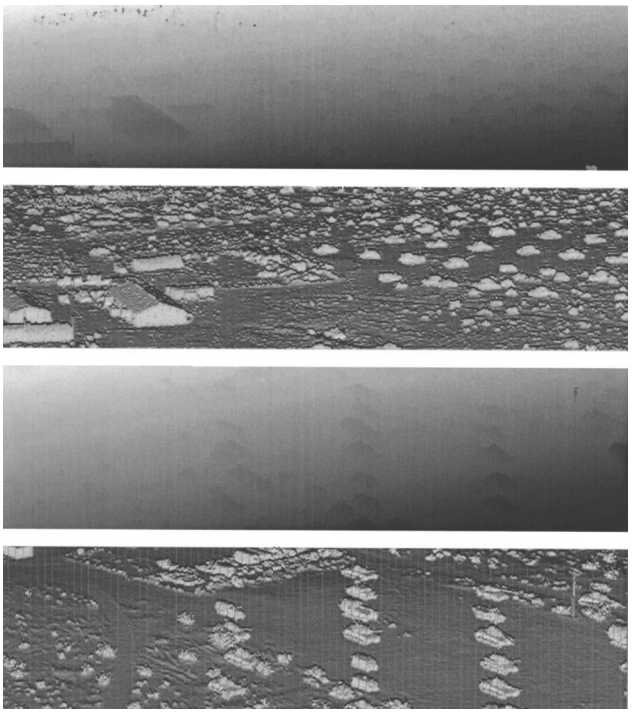

Fig. 12 From top, the first and third images are examples of collected imagery using a ladar seeker, while the second and fourth images are the rendered views.

angle, while metadata provided an estimate of the target aspect angle. While the metadata used may not be practical—or even available—for a tactical ATR system for use in the field, the information serves to limit the scope of the problem and allow for meaningful comparisons.

Given estimates for range, aspect, and elevation, a training set for the test image and its corresponding target chips is developed. Five range bins are used, between ± 50 m of the mean range. Three elevation bins ± 1 deg of the elevation angle estimate are used, while the range in aspect is ± 15 deg with seven bins. This results in 105 training images to be used in building the MACH filters. For the comparison, the M60 model was used to generate the training images, since there are multiple M60 tanks in the test data. All other vehicles would be considered as ''other class'' when computing the Fisher ratio. The same filter parameter list used in the synthetic data section is used again here.

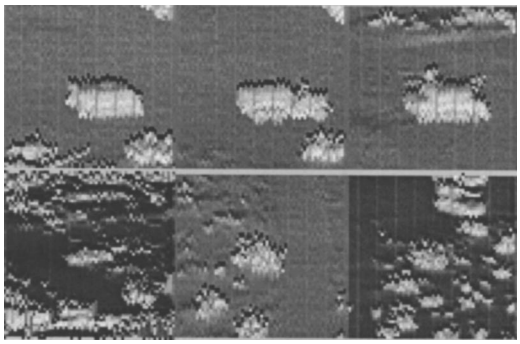

Fig. 13 Examples of confuser targets (top row) and natural clutter chips, which pass some initial detection algorithms (bottom row).

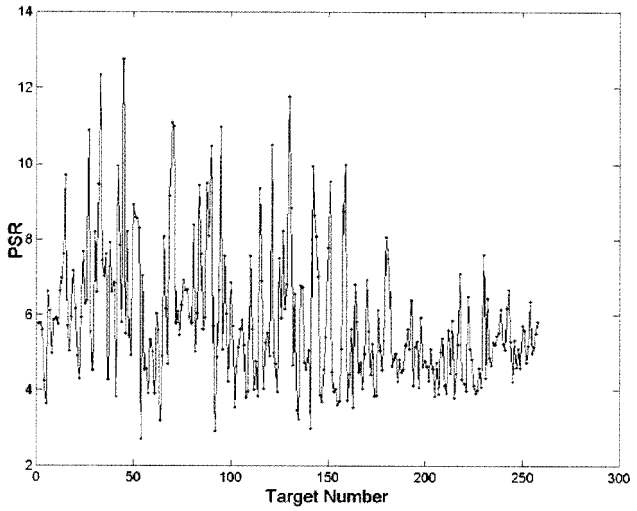

Fig. 14 Plot of PSR versus target number for broadside targets. M60 target numbers end at 141, target numbers >141 indicate nontarget or clutter. Filter parameter set 15 from Table 2 was used to generate the filters.

To generate the PSR data for comparing the filter types, each test chip is filtered using each of the three filters. A plot of PSR versus target number for collection pass is shown in Fig. 14. From this data, over an entire data collection pass ($\sim$10 images, 225 targets, +50 natural clutter objects) the Fisher ratio is calculated. Plots of the Fisher ratio versus the filter parameter number are shown in Figs. 15 and 16.

4 Conclusions

As the plots of PSR indicate, the EMACH filter only outperforms the MACH and GMACH filters when one considers the strict MACH filter, e.g., $\alpha=1$, $\beta=0$, and $\gamma=1$. When only using the ASM term, the EMACH filter does provide a higher average PSR than either EMACH or GMACH. However, in contrast to the results presented in

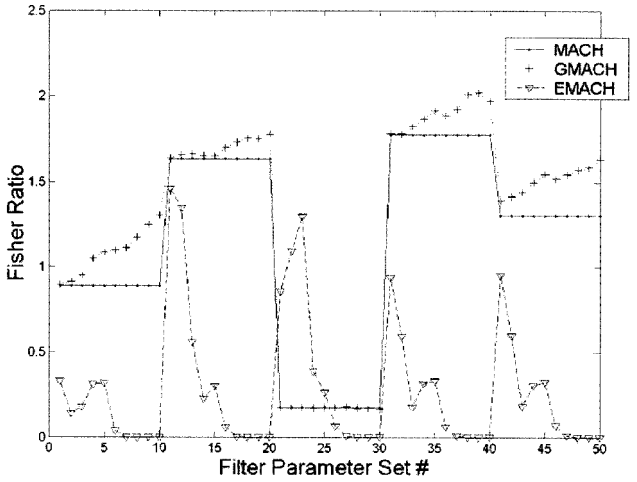

Fig. 15 Plot of Fisher ratio versus filter parameter set number for collected dataset 220 [see Figs. 12(a) and 12(b)]. True class was the M60, while all other vehicles and clutter were considered other class.

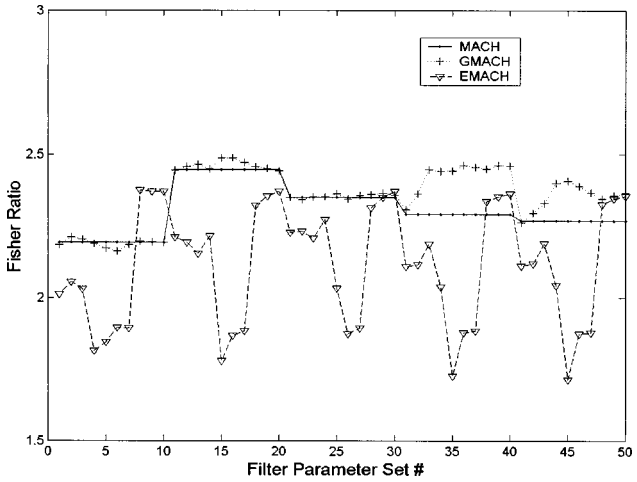

Fig. 16 Plot of Fisher ratio versus filter parameter set number for collected dataset 236 [see Figs. 12(c) and 12(d)]. True class was the M60, while all other vehicles and clutter were considered other class.

Ref. 10, the Fisher ratios for the GMACH and MACH filters are larger than the EMACH filter, suggesting that the choice of filter type is dependent on the nature of the ATR problem and the choice of the training set. Inclusion of the ACE term, however, leads to the EMACH filter performing worse than the MACH or GMACH filters, with lower average PSR values and lower Fisher ratios. The GMACH filter, while it does produce a lower average PSR value for both the true class and false class target, also reduces the variance in the PSR values, resulting in a higher Fisher ratio than either the MACH or EMACH filters. These conclusions seem to hold, considering both the synthetic test cases and the collected ladar test data.

The GMACH filter incurs a higher computational cost than the MACH filter, but this issue is only a concern if one is attempting to build new GMACH filters ''on the fly'' by using estimated target parameters to create a closely matched GMACH filter. In this case, the small decrease in performance of the MACH filters may be acceptable if real-time filter creation is desired.

The inclusion of the new AICH and MASM measures for the EMACH filters provide better performance when one limits the comparison to a narrow class of MACH-type filters. When the comparison is widened to include GMACH and OTSDFs, the data examined here suggest that EMACH will not compare as favorably as suggested in Ref. 10.

Acknowledgments

This work was supported in part by the Office of Naval Research and the Naval Air Warfare Center, Weapons Division, China Lake California.

References

1. D. Carlson, ''Optimal tradeoff composite correlation filters,'' PhD thesis, Carnegie Mellon Univ. (Oct. 1996).
2. A. Mahalanobis et al., ''Unconstrained correlation filters,'' *Appl. Opt.* **33**, 3751–3759 (1994).
3. A. Mahalanobis, B. V. K. Vijaya Kumar, and S. R. F. Sims, ''Distance classifier correlation filters for multiclass target recognition,'' *Appl. Opt.* **35**, 3127–3133 (1996).
4. B. V. K. Vijaya Kumar, D. Carlson, and A. Mahalanobis, ''Optimal tradeoff synthetic discriminant function (OTSDF) filters for arbitrary devices,'' *Opt. Lett.* **19**, 1556–1558 (1994).
5. P. Refregier, ''Optimal tradeoff filters for noise robustness, sharpness of the correlation peak and Horner efficiency,'' *Opt. Lett.* **16**, 829–831 (1991).
6. Z. Bahri and B. V. K. Vijaya Kumar, ''Generalized synthetic discriminant functions,'' *J. Opt. Soc. Am. A* **5**, 562–571 (1988).
7. B. V. K. Vijaya Kumar, ''Minimum variance synthetic discriminant functions,'' *J. Opt. Soc. Am. A* **3**, 1579–1584 (1986).
8. A. Mahalanobis, B. V. K. Vijaya Kumar, and D. Casasent, ''Minimum average correlation energy filters,'' *Appl. Opt.* **26**, 3633–3640 (1987).
9. A. Mahalanobis, D. Carlson, and B. V. K. Vijaya Kumar, ''Evaluation of MACH and Dccf Filters for SAR ATR using the MSTAR public database,'' *Proc. SPIE* **3370**, 460–468 (1998).
10. M. Alkanhal, B. V. K. Vijaya Kumar, and A. Mahalanobis, ''Improving the false alarm capabilities of the maximum average correlation height correlation filter,'' *Opt. Eng.* **39**(5), 1133–1141 (May 2000).
11. M. Alkanhal and B. V. K. Vijaya Kumar, ''Eigen-extended maximum average correlation height (EEMACH) filters for automatic target recognition,'' *Proc. SPIE* **4379**, 424–431 (Apr. 2001).
12. R. Duda and P. Hart, *Pattern Classification and Scene Analysis*, John Wiley, New York (1973).
13. T. Kailath, *Linear Systems*, Prentice Hall, Englewood Cliffs, NJ (1980).
14. D. Duarte and S. Yin, ''Illumination invariant face recognition using composite filters synthesized by simulated annealing,'' *Opt. Eng.* **39**(5), 1252–1259 (May 2000).
15. S. Chang, M. Rioux, and J. Domey, ''Face recognition with range images and intensity images,'' *Opt. Eng.* **36**(4), 1106–1112 (Apr. 1997).
16. A. Van Nevel, C. Kenney, and L. Peterson, ''Image processing for LADAR ATR,'' *Proc. AIAA Missile Sciences Conf.*, AIAA (Nov. 2000).
17. T. D. Cook, ''Results of LADAR ATR captive flight testing experiments,'' *Proc. SPIE* **4379**, 78–85 (Apr. 2001).

Alan Van Nevel is a Research Physicist at the Naval Air Warfare Center, Weapons Division in China Lake, California, and is currently the head of the Image and Signal Processing Branch in the research department. His current interests are in pattern recognition, image processing, target recognition, and the application of physical principles to image processing. In 1990, he completed his BS degree (with honors) at Truman State University, going on to the University of Missouri-Columbia, completing his MS and PhD in physics in 1993 and 1996.

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 interest are in multisensor automatic target recognition, pattern recognition, and image processing. He has more than 90 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 US 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 the 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 at the University of Arizona.