Optical Flow of Small Objects Using Wavelets, Bootstrap Methods, and Synthetic Discriminant Filters

Gary Hewer, Alan Van Nevel, Wei Kuo, and Larry Peterson
Naval Air Warfare Center, China Lake, CA 93555
gary_hewer@mlngw.chinalake.navy.mil
alan@peewee.chinalake.navy.mil
kuo@h3912.chinalake.navy.mil
peterson@luc.chinalake.navy.mil

Abstract

In the look-down/shoot-down scenario, the next generation of air-to-surface missiles will rely on IR sensors and advanced signal processing to detect small (or point) targets in highly cluttered and noisy environments. In this paper, we present a novel wavelet detection algorithm which incorporates adaptive CFAR detection statistics using the bootstrap method. Following detection, the estimate of interframe optical flow is made using synthetic discriminant filters (SDF's). The detection coupled with the new optical flow estimate will enable higher performance in tracking small maneuverable targets. Results for the wavelet bootstrap detection are presented and compared to a conventional matched filter.

1. Introduction

Detection of small targets in the look-down shoot-down situation is becoming increasingly more difficult as target signatures become more noise-like and engagements take place in cluttered environments where the clutter is structured or extended. Such cluttered, lo / signal to noise ratio (SNR) environments push the limits of conventional detection and tracking algorithms. This scenario represents a challenging and unsolved problem. Modern staring focal plane arrays (FPA's) with their superior acquisition ranges and countermeasure resistance can potentially be used to solve the look-down shoot-down problem if two problems can be overcome. First, the structured clutter present in these situations is not fully described by the simple (Gaussian) models commonly used in matched filtering [1] which results in lower detection rates and more false alarms. The second problem is the presence of fixed pattern noise in the IR sensor itself. Fixed pattern noise results from the nonuniform response of the detectors to a uniform incoming signal. Wavelet based image processing enables one to overcome these two problems.

This research has quantified the utility of wavelet-based image-processing algorithms by addressing the point target-detection problem in fixed pattern noise and in structured clutter. This research follows and parallels the standard detection paradigm that has been successfully used in radar detector design. Our research has proceeded by first solving the detection problem when only array noise is present by characterizing the detector array nonuniformity response. Next, we formulated an optimum detection algorithm while recognizing the array noise limits by defining an adaptive constant false alarm rate (CFAR) point target-detection algorithm that maximizes the detection signal-to-clutter ratio. The effectiveness of these detection algorithms are evaluated by using dynamically controlled point targets embedded in a selected set of measured IR backgrounds. The point targets are modeled using a simulated Gaussian target with parametrically varying amplitude, size, and polarity embedded in both fixed pattern noise and in scene-based video images.

Wavelets have properties that suit them to look-down-shoot down antiair and air-to-surface problems; for example, they are scaleable—image features at one scale can be effectively rejected, while other features (like a small target) can be searched out preferentially. Wavelets can extract spatial information (e.g., edges) and can "whiten" the fixed pattern 1/f-type noise present on InSb and HgCdTe staring arrays [2-5]. The number of additions and multiplication's needed to compute the fastest (Daubechies) wavelet transform is directly proportional to the number of pixels in an array and can be implemented on commercial off-the-shelf (COTS) digital signal-processing hardware.

The rest of the paper is organized as follows. The first section gives a brief overview of wavelets, and highlights the reasons that wavelets are beneficial in this case. The section is written to give a general description of wavelet theory, and the reader is encouraged to check the references [6-8] for more details. The second section describes the wavelet CFAR detection algorithm. The algorithm employs bootstrap statistical processing, and this topic is discussed at length in the second section. Also in this section one of the problems

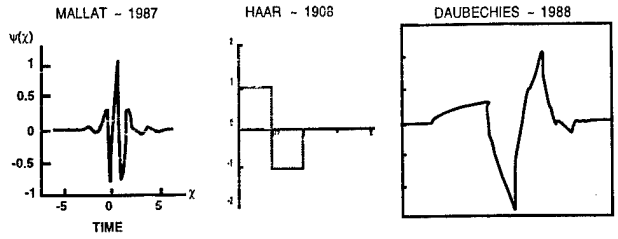

Figure 1 Examples of Wavelets.

discussed in the introduction, namely fixed pattern noise is covered. In addition to the wavelet CFAR detector, the synthetic discriminant filters as developed by Mahalanobis *et al*[9-13] are discussed and how they might be used for optical flow estimates. The third section presents the overview of integrated detection and tracking.. Finally we present some results and conclusions.

2. Wavelet Overview

The development of wavelets is fairly recent, most of the basic theory having been developed in the past 10 or more years with major contributions from researchers in France and the United States. Virtually any function—subject to some simple constraints—can be a wavelet. Since the formalism of the wavelet transform was first introduced by Grossman and Morlet [14], many excellent books [6-8,15-20] including the classic SIAM monograph by Daubechies have appeared. With this wealth of material available, here we will only give a brief overview of wavelet analysis and how wavelets can be brought to bear on the detection problem.

Wavelet transforms translate and dilate a suitably chosen mother wavelet, which decomposes the signal into its local multiscale resolution (coarse to fine). Although wavelets are at least as fundamental as Fourier analysis, they are more flexible and provide information unavailable from the Fourier transform. Like Fourier analysis, wavelets can be interpreted as a basis set in some normed function space. Wavelets are compactly supported (non-zero on some finite interval which provides good localization properties. Localized processing means targets and clutter retain their proper locations before and after processing, which is extremely important for detection. The research of Carmona [21] and Mallat [22] suggests that transients will be localized in certain frequency subbands of the wavelet transform.

An example of a simple wavelet is the Haar wavelet which has compact support, is quite simple, and is defined as follows:

$$\psi(t)=\begin{cases}+1 & 0\le t<1/2\\ -1 & 1/2\le t<1\\ 0 & \text{otherwise}\end{cases}$$

and the companion scaling function is

$$\phi(t)=\begin{cases}+1 & 0\le t<1\\ 0 & \text{otherwise}\end{cases}$$

The panels in Figure 1 illustrate the Mallat, Haar, and Daubechies wavelets, respectively. Although Haar wavelets are simple and compact in the time domain, they are not well localized in Fourier space where they decay very slowly as ω^{-1}, because of the sharp discontinuity. Figure 2 is an example of the standard wavelet transform in a packet architecture. The labels L (scaling coefficients) and H (wavelet coefficients) suggest their role as low- and high- pass filters respectively. As an example, for the Haar wavelet L is an averaging filter and H is a differencing filter. Their elements are $L=\{1/\sqrt{2}\ ,\ 1/\sqrt{2}\}$ and $H=\{1/\sqrt{2},\ -1/\sqrt{2}\}$. The wavelet packet transform is a generalized version of the wavelet transform (Reference **v**). Repeating the wavelet transform by using the output of either or both the low (L)- and high (H)-pass filters as new inputs creates a variable bandwidth filter that is multiresolutional. The input signal is decomposed into low- and high-frequency bands by convolving the signal with the filter and then subsampling (e.g. downsampling by 2).

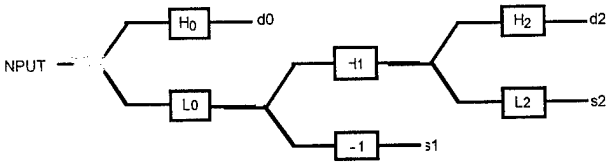

FIGURE 2. Wavelet Packet Architecture .

The two-dimensional transform for images is easily constructed from a tensor product of the one-dimensional wavelet filters in the appropriate order as horizontal and vertical convolutions. As the image is decomposed by these filters, their frequency range divides the image into frequency subbands. Wavelet analysis offers a time-frequency or space-frequency tradeoff that captures low-frequency signals with great frequency accuracy and high-frequency signals with great temporal accuracy—a very reasonable tradeoff, since we can not simultaneously have both time and frequency accuracy , because of the Heisenberg uncertainty principle. The dimension of the wavelet subimages is directly

related to the original image. When the maximally decimated transforms are downsampled by two, the dimension of the image is decreased by 1/2 at each node (i.e., image dimension is decreased by 1/4). Thus, the transformed image will not require any more memory

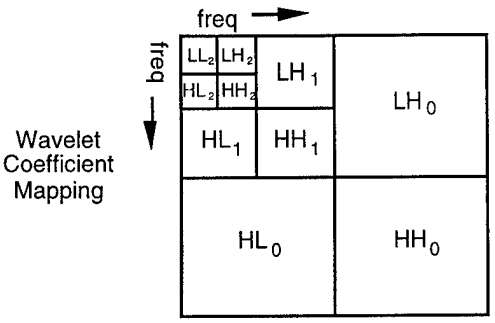

Figure 3. Wavelet Transform

than the original image. Maximally decimated means that the remaining sample after decimation obeys the Nyquist sampling rate.

Figure 3 illustrates a wavelet decomposition of an image into unequal subbands. The ordering of the tree branches are shown in the pyramid structure. The image is split by iterating the four-band decomposition three times. The lowest-frequency components are located in the upper left corner.

In Figure 4 a 512 by 512 image and its wavelet transform are displayed. Smooth regions of the original image will appear in the low-frequency subimage, and sharp edges will appear in the high-frequency subimages. The high pass filters are labeled by d an ' the low pass filters are labeled by s. For example, the d1-d1 256 by 256 subimage in the upper right corner represents the vertical and horizontal convolutions using the H filter. It has mostly all high-frequency features. The 256 by 256 s1-s1 subimage, which has mostly all low-frequency features, represents the vertical and horizontal convolutions with the L filter. It was used as the new input image for the next application of the wavelet transform. The s3-s3 128 by 128 subimage in the lower left corner is the output of the third application of the vertical and horizontal convolutions with the L filter. The white patches are the multiresolution version of the original vehicles; some detail in the subimages is washed out because of the graphical transcription into this document format.

To summarize, wavelets are suitably defined functions that can be used to decompose signals into a time-frequency or space-frequency space. The ability of a few wavelet coefficients to efficiently represent an image within the reduced dimensional subimages means that the computational burdens associated with pattern-recognition algorithms can be accommodated. In addition to being computationally efficient and being perfectly invertible, wavelet transforms have a variety of other properties which make them well suited to the task of point target detection. These properties are described in the next section.

3. Bootstrap, Wavelets and CFAR Detectors

Detection of pulses in the presence of interference or noise requires some type of statistical detection strategy. The wavelet-based detection strategy used in this report is based on the research of Carmona [21]. The functional diagram of our wavelet-based constant false alarm rate (CFAR) detector is shown in the schematic in Figure 5.

A CFAR detector is necessary to control the false alarm rates that are needed to adapt to varying interference and noise. Moreover, any properly designed CFAR detector must constantly monitor and adaptively estimate the CFAR threshold, because a well-known way to defeat a CFAR detector is to raise the noise power gradually [23]. To do this a CFAR processor must adaptively estimate the probability of false alarm rate. Thus, any useful CFAR detector must develop an adaptive test statistic for the false alarm probability P_{fa} = Pr{decide *s(t)* is present : *s(t)* is absent} and the probability of detection P_d = Pr{decide *s(t)* is present : *s(t)* is present}.

The original image is in the upper left hand panel, while it wavelet transform is in the upper right hand panel. After the wavelet transform the pixel values in the subimage are lexicographically ordered by converting the subimage to an n x 1 vector. Modeling the density functions that are directly dependent on the d-scale of the wavelet transform, while perhaps not hopeless, is nontrivial. [24] One fundamental result is known: namely, when the input noise is Gaussian, the density functions of the d-scales are also Gaussian [22] and the variance is decreased by 2^{-j} as the scale index integer j is decremented For this reason the bootstrap technique seems ideally suited to the task of adaptively setting CFAR in complex scenes with non-Gaussian background noise.

The lower left hand panel represents the bootstrap option. A basic principle of the bootstrap method as enunciated by Efron, who named it and demonstrated its scope, is that it represents "the substitution of computational power for theoretical analysis. The payoff, of course, is freedom from the constraints of traditional parametric theory, with its over reliance on a small set of standard models for which theoretical solution are

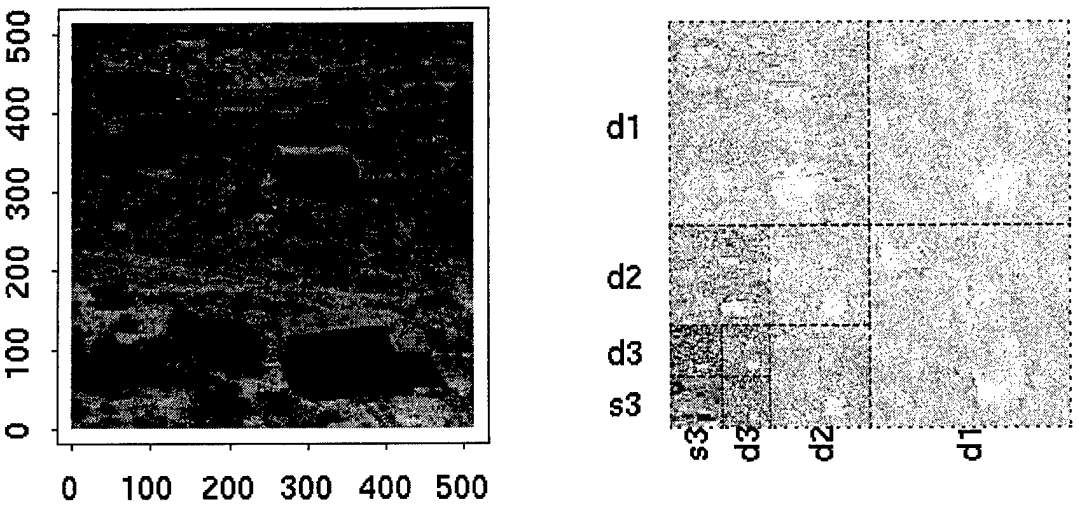

Figure 4. Original Image and Wavelet Subimages.

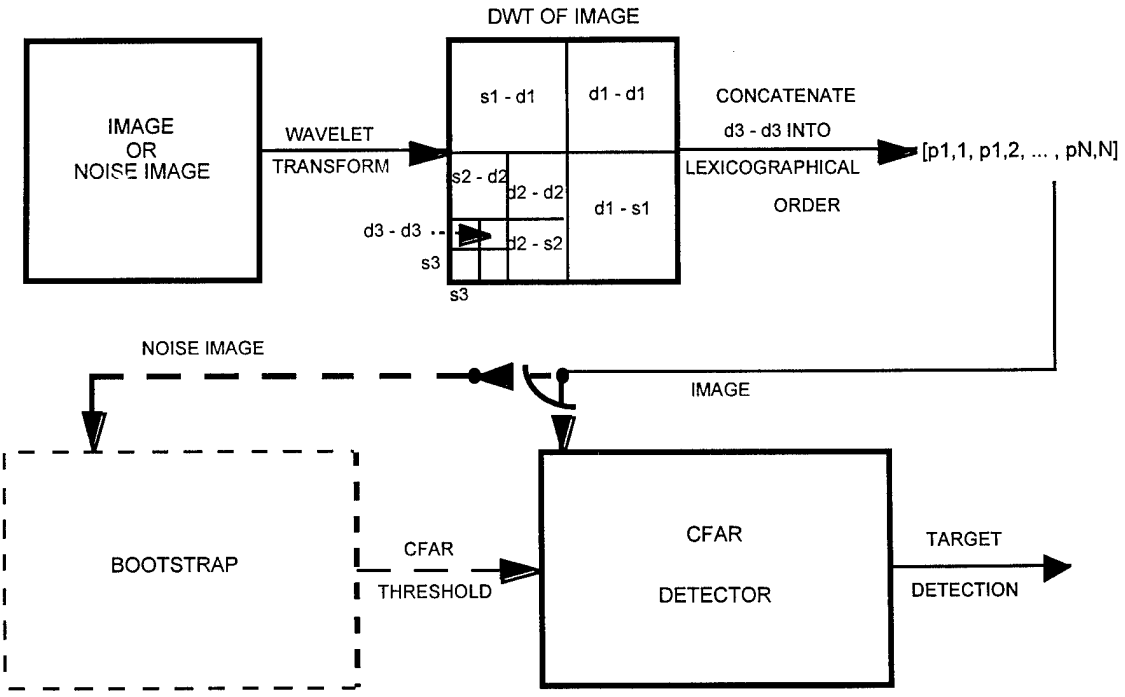

Figure 5. Wavelet-Based CFAR Detector.

available" [25]. Thus, the bootstrap relieves us from having to make parametric assumptions about the underlying background noise. However, it does have its own modeling requirements such as the degree of correlation present in the wavelet coefficients and the required number of training sets for valid inference. According to Hall, the insight of Efron was to realize that in complex situations, when bootstrap statistics are awkward to compute, they may be approximated by Monte Carlo "resampling" [26]. The name "bootstrap" refers to the use of the original data set to generate new data sets by

sampling with replacement. A schematic of the basis bootstrap resampling processing is shown in Figure 6.

Given a sample x* of size n, random samples of size n are drawn with replacement B times from the original sample, and the value of a statistic is then computed for each sample. Although there is not a complete and rigorous theory for our extension of the bootstrap to these problems, we take our inspiration from the following guidelines found in the paper Efron and Tibshirani [27], namely, (1) the bootstrap algorithm can be applied to almost any statistical estimation problem or data structure, and (2) "the statistic t(**x**) can be anything at all, as long as we can compute t(**x***) for every bootstrap data set **x***."

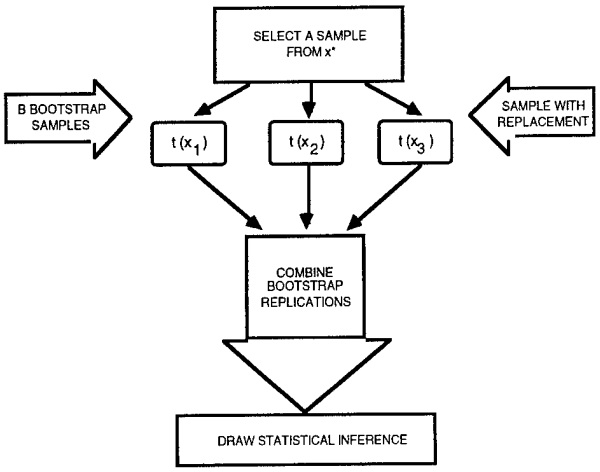

FIGURE 6. Basis Bootstrap Resampling Process.

The computational overhead associated with the bootstrap is directly proportional to B times the number of operations in the computation of T(d*). For a sample of size n, the number of operations is proportional to B*O(n) for a maximally decimated Daubechies wavelet transform and to B*O(n log n) for an undecimated transform.

3.1 MATCHED FILTERING

It is well known that the optimal detector for known signals in additive white Gaussian noise is the matched filter, which is normally implemented by correlating the received signal, $r(t)$, with the known signal, $s(t)$. In discrete time we compute

$$g(\mathbf{r}) = \mathbf{r}^T \mathbf{s}$$

where **r** and **s** are vectors containing samples of the received signal and known signal. These statistics are compared to a threshold to decide whether the signal **s** was present during the observation interval. When the arrival time of the signal is unknown one typically correlates at each possible time shift. The output of the filter exceeding a threshold indicates the presence of the signal, and the location of the peak output indicates the location of the signal.

For non-white Gaussian noise the optimal discrete-time detector uses essentially the matched filter modified by multiplication by the inverse of the noise covariance matrix:

$$g(\mathbf{r}) = \mathbf{r}^T \boldsymbol{\Sigma}^{-1} \mathbf{s}$$

Another well-known result in detection theory is that, for known signals in additive Gaussian noise, one can expand the known signal and the received signal in another orthonormal basis and perform the detection process in the coordinate system without degrading the overall performance. Daubechies' wavelets are an orthogonal transform. When the known signal is corrupted by non-Gaussian noise, there is usually no simple derivation for the optimal detector. In many cases the optimal detector is unknown, which explains the existence of numerous competing suboptimal approaches.

To evaluate the performance of the wavelet bootstrap detector it is necessary to embed and compare it within the matched filter context. Figure 7 is a block diagram illustrating the details of matched filter detection. The original derivation is by Chen and Reed [28] with helpful discussions in [29]

Before we explain the concepts associated with the main block s, some of the terms should be defined. An image chip is a sub-block of pixels on the array; often it will be a 3 by 3 or 4 by 4 square of pixels. The noise sample is another image chip, ideally devoid of targets, that is selected to estimate the covariance of the background noise/clutter. The template s is a model of the target and can be quite detailed. For small targets it is often a simple numerical approximation for the optical point spread function. Using CFAR detectors is equivalent to testing the hypothesis H_0 (no target present, only clutter/noise) denoted by $\mathcal{E}$ versus the alternative hypothesis H_1 (target and clutter/noise present) denoted by $b + \mathcal{E}$ by using a simple threshold. The clutter covariance is denoted by $\Sigma_{\mathcal{E}}$, and its inverse multiplied by the transpose of the lexicographically ordered template is the matched filter. If the clutter distribution is Gaussian and the filter is linear, then the filtered clutter is also Gaussian. Assuming a Gaussian clutter model, the Gaussian log likelihood ratio test [28] is designed to make the optimum distinction between the two hypotheses. The threshold is set by noting that these assumptions are equivalent to testing a mean shift mod-

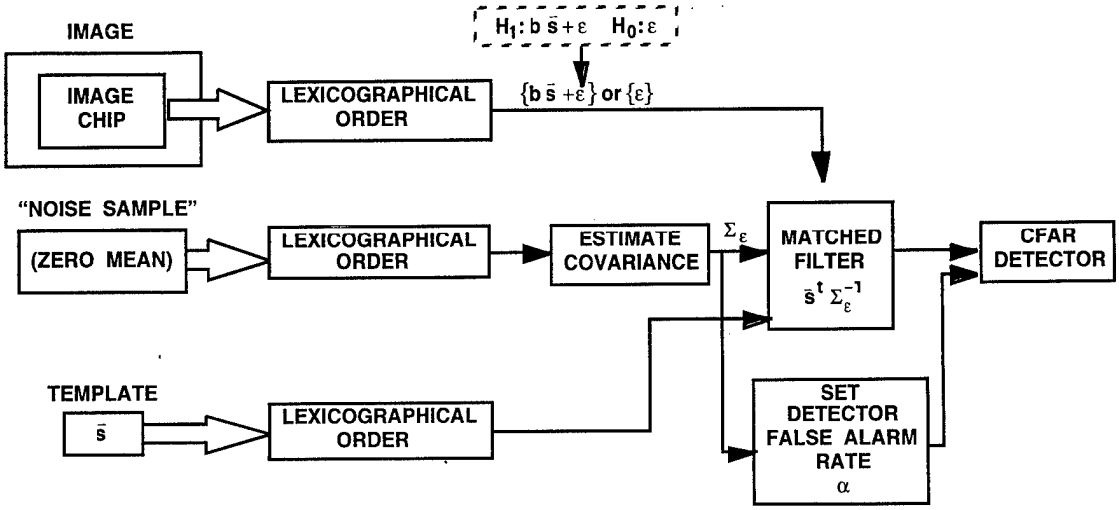

Figure 7. Matched Filter Detection.

eled by the template s in the multivariate normal distributions

$$\mathbf{H}_1\ \mathbf{x} = \mathbf{s} + \varepsilon \quad x \approx \mathrm{N}(\mathrm{s}, \Sigma)$$

$$\mathbf{H}_0\ \mathbf{x} = \varepsilon \qquad x \approx \mathrm{N}(0, \Sigma)$$

Thus, the false alarm can be estimated using the multivariate normal with zero mean and sample variance $s^t \Sigma_\varepsilon^{-1} s$. Setting the CFAR rate is a critical step for infrared images, because if the background clutter is not always well approximated by the Gaussian hypothesis, higher false alarm rates will result.

3.2 Spatial Clutter Statistics

If the Gaussian assumption is violated, outliers will be present and manifest themselves in the tail of the distribution. Outliers are "rogue observations" that lie, in some sense, far from the middle of the data and bias the false alarm estimate. A quick and effective means with which to visualize the tail behavior in empirical data is provided by the quantile-quantile plot (Q-Q plot). It can be used to compare the degree of agreement between two empirical distributions, or it can be used to compare the empirical quantiles with the quantiles from an ideal distribution. Two distributions are compared by graphing quantiles of one distribution against the corresponding quantiles of the other. A normal Q-Q plot is a plot of the ordered data y_i from the sample $\{x_0, \cdots, x_{N-1}\}$ versus $y_{p_i} = \Phi^{-1}(p_i)$, where $p_i = (i - 1/2)/\mathrm{N}$, $\mathrm{i} = 1, 2, \cdots, \mathrm{N}$, and Φ^{-1} is the inverse of the standard normal distribution [30]. If the shape of the unknown distribution is approximately normal, even in the tails, then the empirical quantile sample values will approximate the normal line. A normal Q-Q plot will reveal the presence of outliers in the extreme tails or leptokurtic shape. Both are strong indicators that robust methods are required. The shape is leptokurtic as defined in Cleveland [31]*if the relative density of the data in the middle of the distribution compared with the density in the tails is thinner than the normal distribution.* Figure 8 displays a Q-Q plot using a 64 by 64 subimage taken from a Skyball image. These samples clearly exhibit non-normal behavior in the tails. Moreover, the subimage dimensions are typical of the subimage tiling that is often used to localize the scene statistics.

3.21 Fixed pattern noise

In a series of seminal papers, Flandrin [32], and Wornell [33] proved that the Daubechies wavelets will decorrelate or "whiten" a broad class of stochastic processes that have a 1/f-type spectra. Earlier, Scribner and others [2-4] have observed that a major contributor to the fixed pattern nonuniformity observed on infrared staring arrays is primarily due to the 1/f-type spectra associated with the sampled output of the individual pixels. Using measurement from an Amber 128 by 128 staring array, Hewer and Kuo [5] combined these two results to demonstrate that the Daubechies wavelets do "whiten" the 1/f-type spectra as predicted. These transforms effectively decorrelate the fixed pattern noise inherent on indium antimonide (InSb) and mercury cadmium telluride (HgCdTe) starring arrays.

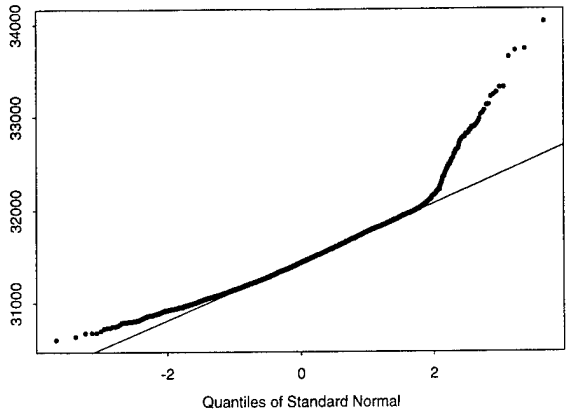

Figure 8. Q-Q plot of data from a Skyball image. Note the deviation from the straight line in the tails, indicating non normal behavior.

3.3 Optimal trade-off synthetic discriminant filters (SDF's)

As an adjunct to the matched filtering, a class of correlation filters as developed by Mahalanobis *et al* [9-13], has been implemented. These filters have exceptional tolerance to scaling and rotation distortions. The tolerance of the filters is incorporated through the selection of an appropriate training set, and can be tuned to provide high (generalization) or low (specificity) tolerance.

In the discussion of the MACH (maximum average correlation height) filters that follows, bold lowercase indicates a column vector, while bold uppercase represents a diagonal matrix. The filters result from maximizing the ratio

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

where **h** is the correlation filter and **m** is the average of the training images in the Fourier domain. Each image is lexigraphically ordered to form a vector. **S** is the average similarity measure matrix

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

In eq. (2) $\mathbf{X}_k$ are the individual training images, again in the Fourier domain. The training image is lexigraphically ordered and its elements placed on the diagonal of $\mathbf{X}_k$. **M** is the mean training image, arranged similarly to $\mathbf{X}_k$. 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 **h** is then given by

$$\mathbf{h} = \mathbf{S}^{-1}\,\mathbf{m} \qquad (3).$$

Variants on the MACH filter can be achieved by varying the performance metric one wishes to maximize. For example Refrieger [13] has developed optimal trade-off synthetic discriminant filters (OTSDF's) which attempt to minimize the energy functional

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

where

$$\mathbf{Q} = \alpha\,\mathbf{P} + \beta\,\mathbf{D} + \gamma\,\mathbf{S} \qquad (5).$$

S is as defined previously, while **P** is the power spectral density of the expected noise, and **D** is the average power spectral density of the training set. The constants $\alpha, \beta, \gamma, \delta$ are non-negative and must satisfy $\alpha^2 + \beta^2 + \gamma^2 + \delta^2 = k$ where k is any positive constant. Minimizing E(**h**) results in

$$\mathbf{h} = \frac{\delta}{2}\mathbf{Q}^{-1}\mathbf{m} \qquad (6).$$

By varying the parameters, one can adapt the filter for the optimal performance for the situation under study. If one sets $\alpha = \beta = 0$, the result is the 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 (DCCF's), which are able to distinguish between multiple classes of similar objects (*e.g.* T72's vs. M1A1 tanks).

The class of OTSDF filters was chosen for the feature detection for several reasons. As discussed, the filters can incorporate varying degrees of distortion tolerance and be built to generalize classes of targets. Another benefit of the algorithm is that the result is statistically optimum, and depends on a realistic, mathematically rigorous optimaztion 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. However, the limit of its performance may be reached when considering small (< 5x5 pixels).

This limit is to be investigated more thoroughly in future work.

4. Integrated detection and tracking

The functional diagram of our overall integrated detection and tracking design is shown in the schematic in Figure 9. The original image is in the upper left hand panel, while it wavelet transform is in the upper right hand panel. The two images in the lower left hand panel illustrate optical flow. Optical-flow fields, which describe image domain motion, extracted from two or more images, must discriminate between motion and local illumination changes due to the intrinsic noise field on the array, and the illumination changes due to the clutter field in the sensor's field of view. Being able to accurately estimate the velocity field when sub-pixel and multi-pixel interframe changes are present is essential and difficult especially for dim targets moving relative to the background . Nobody has a satisfactory solution for optical flow , when dynamic occlusion is present—the estimation of optical flow is not yet a mature subject. The highlighted streak in the lower right hand image represents a track history from a Kalman filter. When detecting low intensity targets in dense clutter the allowable false target confirmation rate is a system requirement. Thus, an integrated system must include a system tradeoff between tracking complexity, scan rate, and detection threshold. The tracking methodology in this research will borrow heavily from the research of Blackman [34] and Bar-Shalom [35]. The challenge here is to integrate the wavelet-bootstrap detector and the SCF for multipixel targets with their data association algorithms and Kalman filters for track file maintenance.

The SDF's are used to estimate the interframe motion. Given a set of potential targets as detected by the wavelet matched fi'.er, the SDF detection/classification is performed to yield an accurate assessment of the position of the target. In the next frame, the SDF algorithm is performed again, searching in areas which previously contained targets. From these two frame results, it is possible to estimate the target motion. With the motion estimate in hand, it is possible to feed this data into a Kalman filter and update the filter more accurately. In some test cases, extremely accurate tracking performance was demonstrated. The lower limits of the algorithm, in terms of SNR and target size, is yet to be determined, and is a basis for future work.

5. Results

5.1 Imagery and Data

Before presenting our results, a brief description of the data sets that were used will be presented. Several databases have been used in this effort. Blackbody measurements to characterize fixed pattern noise were obtained during our early research using an Amber AE 4128 128 by 128 InSb staring array that operates at midwave infrared wavelengths. Additional blackbody images from a Honeywell uncooled microbolometer 336 by 165-pixel IR sensor at long wave infrared wavelengths were obtained from D. A. Scribner, J. T. Caulfield, and M. R. Kruer of the Electro-Optical Technology Branch at the Naval Research Laboratory

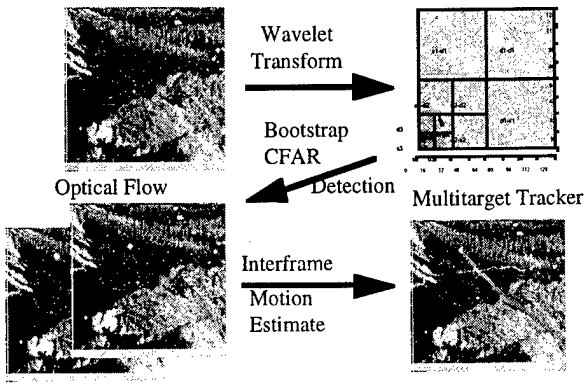

Figure 9 Integrated detection and track

(NRL), Washington, D.C. A bolometer is an infrared detector that measures absorbed, incident infrared radiation by a voltage change in electrical resistance due to a nonequilibrium temperature differential. They also provid[illegible]0 Amber 256 by 256 images taken during tests overlooking the Chesapeake Bay.

Additional digital IR image sequences collected from an airborne platform using the Skyball two-color IR seeker. Skyball was a joint, 1.5-year, digital data-collection effort by the Naval Air Warfare Center Weapons Division, China Lake, California, and Hughes Missile Systems. The Skyball sensors consisted of two 128 by 128 HgCdTe FPAs operating simultaneously in both mid- and long-wave IR spectral bands mounted in the nose of a Jetstar-8 aircraft with an in-flight blackbody calibrator. Typical frame rates varied typically from 25.6 to 49.1 frames per second. Unclassified Skyball image sets include a statistically significant sample of cluttered backgrounds flown at varying ranges, geometries, and seasons. Very large format, extremely high spatial resolution 12-bit digital IR image sequences were also acquired from the Airborne Infrared Measurement System program sponsored by

DARPA. A full description of the data can be found in reference [36].

To quantify algorithm performance, an interactive approach between algorithm testing and development was employed by using synthetic targets. Routines were written to generate subpixel-size to N- by N-size raised Gaussian targets. Target sizes are presently limited to subpixel up to 7 by 7 pixels. This range has been adequate for small target-detection investigations to date. The method can be readily extended to larger sizes, should this be required. The software enables the researcher to insert anti-aliased Gaussian targets at user-specified pixel locations with user-designated peak amplitudes. Polarity reversal of contrast is supported so that targets may fall through the neighboring background level and rise back up through it. Peak amplitude, amplitude change, rate of change, size, and location are specified by the user and can be adjusted to meet requirements. Care has been taken during the rendering process to preserve target edge contrast and minimize aliasing so that any artificial edge enhancement due to target insertion is minimized

5.2 Results

All of the comparative detectors were applied to the d3-d3 subimage described in Figure 3. For very small targets occupying a square image chip (i.e., <5 X 5 or fewer pixels), the matched filter shape is matched to the blur circle. The three matched filter detectors using a sampling annulus are: 1) empirical covariance assuming a Gaussian distribution to set the detection threshold [1], 2) empirical covariance using a bootstrap method to set the detection threshold, 3) autocovariance using a bootstrap method to set the detection threshold. Two other detectors, namely, 4) a matched filter without a sampling annulus combined with an autocovariance and bootstrap threshold 5) a bootstrap threshold extracted directly in the d3-d3 subimage were also studied. statistics. All video sequences were fifty frames in length and used embedded Gaussian stationary targets. The CFAR detection statistics were estimated from the d3-d3 subimage of the first frame and then applied to the entire sequence. Further details are forthcoming in the NAWCWPNS Technical Publication [39]. If the sampling chip in Figure 9 should accidentally include the target, then a lower-than-expected probability of detection will result [27]. To avoid this possibility a sampling annulus, as suggested by Singer and Saski [1], is placed around the suspected target area and the covariance estimate is obtained from it

The background clutter used in these investigations included Skyball and Chesapeake Bay IR imagery. The Skyball sequences selected for the comparative study represent mountainous and urban – light industrial and residential terrain. The mountainous background data collected during daylight hours over the Sierra Nevada is comprised primarily of forest, snow and rock. The mid-wave infrared urban images were collected at night along U.S 101 near the Santa Barbara, CA coast. The sequences included light industrial urban areas with the highway and freeway interchange visible, small to large several story buildings and residential areas. The Chesapeake Bay video imagery comprised of ocean and sky background c was collected using an on shore stationary Amber camera.

The standard SNR as given in [38] is defined as

$$\mathbf{SNR} = 10\log_{10}\left(\frac{\frac{1}{n}\sum_{j=1}^{n} signal^2}{\mathrm{var}(noise)}\right)$$

where $signal$ is pixel amplitude, j denotes pixel number, and n is the number of target pixels in the sampling chip. In these studies a 10 x 10 chip was used.

The design of a simple binary hypothesis detector is such that any pixel value that exceeds the threshold will be classified as a target. If a single detection threshold is used for the entire infrared array, then any patchy outlier clutter response or fixed pattern noise will force a compromise between the false alarm rate and the local target-to-clutter ratio. Thus, the standard receiver operating characteristic or CFAR performance curves that use the dichotomous probability of detection versus probability of false alarm will require very high target-to-clutter ratios. Such high ratios will understate the performance of any detector that has a high probability of detecting the target, while simultaneously admitting an acceptable number of threshold exceedances. One major function of Kalman filtering is to provide data association files to manage threshold exceedances. For this reason, the number of false positives is a useful performance metric.

In the Chesapeake Bay sequence, all detectors were able to detect the target with a probability of detection (Pd) of 100% down to a SNR of -4 dB. However, at the lowest detected SNR the average number of false positives per frame for the fifty frames was quite variable. The range was as follows, Detector 1 has 0.52 false positives, Detector 2 had 2.42 false positives, Detector 3 had 6.8 false positives, Detector 4 had 16.02 false positives, and Detector 5 had 4.02 false positives. This data set was unique in that the dynamic range of the pixel amplitudes in each of these frames was 6-bits.

In the Skyball IR long wavelength mountainous terrain sequence, Detector 1 had a Pd of 6% with 0.33 false positives per frame at a SNR of 2 dB. It could not reliably detect targets at lower SNRs. Detector 2 had a Pd of 6% with 3.33 false positives at a SNR of -5 dB while Detector 3 had Pd of 30% with 1.33 false positives. Detector 4 exhibited 100% Pd down to a SNR of -5 dB with 13.42 false positives per frame, and Detector 5 had a Pd of 6% with 0.67 false positives at –5 dB SNR.

In the Skyball mid-wavelength IR mountainous terrain sequence, Detector 1 had a Pd of 4% with 0.5 false positives at a SNR of -6 dB. It had no further reliable detections at lower SNRs. At a SNR of –10 dB Detector 2 had a Pd of 18% with 2.78 false positives , Detector 3 had a Pd of 18 % with 5.0 false positives, Detector 4 had a Pd of 100% with 10.70 false positives per frame. Detector 5 had Pd of 2% with 2.0 false positives at SNR of -9 dB. It was unable to reliably detect the target below –9dB SNR.

In the Skyball mid wavelength urban-light industrial sequence, Detectors 1, 2, and 5 could not reliably detect the target at a SNR of -1 dB. No higher SNR data were run with this background. Detector 3 had a Pd of 94% with 1.19 false positives at a SNR of -6 dB while Detector 4 had a Pd of 100% with 13.46 false positives. At a SNR of –13 dB Detector 3 had a Pd of 50% with 1.0 false positives while Detector 4 had a Pd of 52% with 13.77 false positives .

5.3 CONCLUSIONS

Wavelet transforms combined with matched filtering represent a computationally efficient algorithm to locate small target candidates on an imaging array. These filters are compatible with CFAR detectors methodology. However, robustly setting the CFAR threshold is still one of the biggest challenges in infrared detection, because of the limitations of parametric models to capture the structured clutter diversity. For this reason the results are open to further interpretation and research. While the bootstrap can accommodate the non-Gaussian clutter distributions, its overall performance will depend on being able to manage the resulting false alarm rate, which ultimately must be evaluated within a track file context in order assess its final impact.

The use of synthetic discriminant filters to estimate optical flow of small targets is a new approach. The limiting factor is the target size. Further work needs to be done do determine the lower limit on target size. It is known that reliable results can be achieved with targets of 5x5 pixels; results for smaller targets are forthcoming. For midcourse corrections and the end game scenario, the SDF's are able to improve the accuracy of the tracking, and also able to discriminate between types of targets.

REFERENCES

[1] P. F. Singer and D. M. Sasaki. "The Heavy-Tailed Distribution of a Common CFAR Detector," in *Proceedings of the Society of Photo-Optical Instrumentation Engineers*, 1995, pp. 124-140. (Also in *Signal and Data Processing of Small Targets*, SPIE Vol. 2561 (1995).)

[2] D. A. Scribner, M. R. Kruer, C. J. Gridley, and K. Sarkady. "Measurement, Characterization, and Modeling of Noise in Staring Focal Plane Arrays," *Infrared Sensors and Sensor Fusion*, SPIE Vol. 782 (1987), pp. 147-160.

[3] _____. "Physical Limitations to Nonuniformity Correction in IR Focal Plane Arrays," *Focal Plane Arrays: Technology and Applications*, SPIE Vol. 865 (1987), pp. 185-202.

[4] D. A. Scribner, K. A. Sarkady, M. R. Kruer, and C. J. Gridley. "Test and Evaluation of Stability in IR Staring Focal Plane Arrays After Nonunifor mity Correction," *Test and Evaluation of Infrared Detectors and Arrays*, SPIE Vol. 1108 (1989), pp. 255-264.

[5] G. A. Hewer and W. Kuo, *Wavelet Transform of Fixed Pattern Noise in Focal Plane Arrays*, NAWCWPNS, February 1994. (NAWCWPNS TP 8185, publication UNCLASSIFIED.)

[6] Ingrid Daubechies. *Ten Lectures on Wavelets.* CBSM-NSF Regional Conference Series in Applied Mathematics, Vol. 61, Society for Industrial and Applied Mathematics, Philadelphia, 1992.

[7] G. Strang and T. Nguyen. *Wavelets and Filter Banks*. Wellesley, MA, Wellesley-Cambridge Press, 1996.

[8] Stephane G. Mallat, "A Wavelet Tour of Signal Processing", New York, Academic Press, 1998

[9] D. Carlson, "Optimal tradeoff composite correlation filters," Ph.D. thesis Carnegie Mellon Univ., Oct. 1996.

[10] A. Mahalanobis *et al*, "Unconstrained correlation filters," *Applied Optics*, vol. 33, pp. 3751-3759, 1994.

[11] A. Mahalanobis, B.V.K. Vijaya Kumar, and S.R.F. Sims, "Distance classifier correlation filters for multiclass target recognition," *Applied Optics*, vol. 35, pp. 3127-3133, 1996.

[12] B.V.K. Vijaya Kumar, D. Carlson, and A. Mahalanobis, "Optimal tradeoff synthetic discriminant function (OTSDF) filters for arbitrary devices," *Optics Letters*, vol. 19, pp. 1556-1558, 1994.

[13] Ph. Refregier, "Optimal tradeoff filters for noise robustness, sharpness of the correlation

peak and Horner efficiency," *Optics Letters*, vol.16, pp. 829-831, 1991.

[14] A. Grossman and J. Morlet. "Decomposition of Hardy Functions Into Square Integrable Wavelets of Constant Shape," *SIAM J. Math, Anal.*, Vol. 15 (1984), pp. 723-36.

[15] R. Carmona, Wen-Liang Hwang, Bruno Torresani, "Practical Time-Frequency Analysis : Gabor and Wavelet Transforms With an Implementation in S (Wavelet Analysis and Its Applications, V. 9)", New York, Academic Press, 1998.

[16] Yves Meyer. *Wavelets: Algorithms and Applications*, tr. and rev. by Robert D. Ryan. Philadelphia, Society for Industrial and Applied Mathematics, Philadelphia, 1993.

[17] M. Vetterli and J. Kovacevic. *Wavelets and Subband Coding*. New York, Prentice-Hall, 1995.

[18] Mladen Victor Wickerhauser. *Adapted Wavelet Analysis From Theory to Software*. Wellesley, MA, A. K. Peters, 1994.

[19] J. M. Combes, A. Grossman, and P. Tchamitchian, eds. *Wavelets, Time-Frequency Methods and Phase Space*. New York, Springer Verlag, 1988.

[20] C.K. Chui, *An Introduction to Wavelets*,San Diego, Academic Press 1992.

[21] R. Carmona. "Wavelet Identification of Transients in Noisy Time Series," *SPIE Proceedings 2034 (1993), Mathematical Imaging: Wavelet Applications in Signal and Image Processing*, Bellingham, WA.

[22] S. G. Mallat and W. L. Hwang. "Singularity Detection and Processing With Wavelets," *IEEE Trans. on Information Theory*, Vol. 38 (1992), pp. 617-43.

[23] G. Minkler and J. Minkler. *CFAR*. Baltimor , MD, Magellan Book Co., 1990.

[24] E. Simoncelli, "Statistical Models for Images: Compression, Restoration and Synthesis," *Proc of the 31st Asilomar Conf. on Signals Systems and Computers*, Nov. 1997.

[25] Bradley Efron. *The Jackknife, the Bootstrap and Other Resampling Plans*. CBSM-NSF Regional Conference Series in Applied Mathematics, Vol. 38, Society for Industrial and Applied Mathematics, Philadelphia, 1982.

[26] Peter Hall. *The Bootstrap and Edgeworth Expansion*. New York, Springer-Verlag, 1992.

[27] Bradley Efron and Robert J. Tibshirani. "Statistical Data Analysis in the Computer Age," *Science*, Vol. 253 (July 1991), pp. 390-395.

[28] J. Y. Chen and I. S. Reed. "A Detection Algorithm for Optical Targets in Clutter," *IEEE Trans. on Aerospace and Electronic Systems*, Vol. 23 (January 1987), pp. 46-59.

[29] M. Scheffe, M. M. Blane, and D. B. Cooper. "Multicovariance Matched Filter for Target Detection and Background Recognition," in *Proceedings of the Society of Photo-Optical Instrumentation Engineers*, 1995, pp. 203-218. (Also in *Signal and Data Processing of Small Targets*, SPIE Vol. 2235 (1994).)

[30] R. Gnanadesikan. *Methods for Statistical Data Analysis of Multivariable Observations*. New York, John Wiley & Sons, 1977.

[31] W. S. Cleveland. *Visualizing Data*. Summit, NJ, Hobart Press, 1993.

[32] P. Flandrin. "On the Spectrum of Fractional Brownian Noise," *IEEE Trans. on Information Theory*, Vol. 36 (1989), pp. 197-199

[33] Gregory W. Wornell. "A Karhunen-Loeve Like Expansion for 1/f Processes via Wavelets," *IEEE Trans. Inform. Theory*, Vol. IT-36 (1990), pp. 859-861.

[34] S. Blackman, *Multiple Target Tracking with Radar Applications*, Artech House, 1986.

[35] Y. Bar-Shalom and X-R Li, *Estimation and Tracking: Principles Techniques and Software*, Artech House, Norwood, MA, 1993.

[36] K. Ralston, J.B. Attlil, R.Fries, C. Acquista, P. Clemes, Paula Henderson, C. Chan, A. Dalcher, P. Wiley, H. Axelrod, P. Bender, B. Sturges, S. Blackman, F. Boyton, and W. Helliwell, " Airborne Infrared Measurement System (AIRMS) Final Technical Report.", DARPA, Dec. 31, 1996.

[37] M. B. Priestly. *Spectral Analysis and Time Series*. Vol. 1, Univariate Series. New York, Academic Press, 1981.

[38] *The Infrared and Electro-Optical Systems Handbook, Vol. 4*. ERIM and SPIE Optical Engineering Press, 1993.

[39] Gary Hewer, Wei Kuo, and Larry Peterson *Wavelet Detection of Small Stationary Targets in Infrared Imagery* , NAWCWPNS TP 8398 To be published.