DETECTION OF POINT TARGETS IN IMAGE SEQUENCES BY HYPOTHESIS TESTING: A TEMPORAL TEST FIRST APPROACH

Alexis P. Tzannes

Air Force Research Laboratory
Hanscom AFB
alexis@max.rl.plh.af.mil

Dana H. Brooks

CDSP Center, ECE Department
Northeastern University
brooks@cdsp.neu.edu

ABSTRACT

This paper addresses the problem of designing an efficient and effective image sequence processing scheme that will successfully detect very small (point) targets in a cluttered background when both the target and clutter are moving through the image scene. The specific application area was detection of targets such as airplanes in infrared (IR) image sequences of a cloudy sky which have been taken by a stationary camera. In general we assume that targets are typically one to two pixels in extent and move only a fraction of a pixel per frame, are often low amplitude, and are found in scenes which also contain evolving clutter, *e.g.* clouds. Our algorithm is based on signal processing and detection theory, includes a perfect measurement performance analysis, and can be made computationally efficient compared to other approaches. Thus the algorithm could be applicable to other image sequence processing scenarios, using other acquisition systems besides IR, such as detection of small moving objects or structures in a biomedical or biological imaging scenario or the detection of satellites, meteors or other celestial bodies in night sky imagery acquired using a telescope. We present a GLRT solution, perfect measurement analysis including ROC curves, and results using real-world infrared data.

1. INTRODUCTION

The problem of detecting and tracking point targets in image sequences has been an active area of research for several years. Initial algorithms separated the problem into a spatial detection stage, followed by a temporal association or tracking stage [1, 2]. These algorithms are adequate when the targets are bright compared to the background but perform poorly with dim, small targets in severe clutter, since such targets may not pass the initial spatial detection and thresholding stage. More recent approaches used multiple frames to incorporate temporal as well as spatial information, often referred to as "track before detect" algorithms [3, 4]. The standard approach poses the tracking problem as the detection of a known signal in 3-D noise. The main drawbacks of these techniques are that they depend on specific statistical noise and clutter distributional models which may not correspond well to real data, and are very computationally intensive since the entire 3-D space must be filtered for all possible trajectories for each target velocity. Suboptimum approaches that use dynamic programming have been proposed to reduce the computational complexity [5, 6] at the cost of reduced performance, especially for dim targets, or assume the clutter can be whitened and use sequential estimation [7].

The original idea of our approach, first proposed in [8, 9], was that temporal processing first (that is, filtering the profile of each pixel in time before spatial processing) can be a powerful tool in detecting point targets, providing good clutter suppression at relatively low computational cost. Temporal processing is effective because it exploits the difference between the temporal profiles of pixels through which a target passes, compared to those affected by clear sky or cloud clutter. Pixels that see clear sky or other features which are constant in time will have time profiles that are white noise. A pixel that is affected by a point target will have a pulse like shape, with the width of the pulse being inversely proportional to the target velocity. Pixels that are affected by cloud edges or other clutter features will have temporal profiles that behave less regularly, but which may be amenable to statistical modeling. Temporal processing is computationally efficient because it involves 1-D processing first, and can eliminate the vast majority of pixels as candidates before 2-D processing is employed.

The novelty in this paper is that it builds on the intuition in [8, 9] by formulating the detection problem in a theoretically sound hypothesis testing framework. In this framework we introduce deterministic and statistical models for targets, clutter, and clear sky which we have developed through the study of a large database of real IR sequences, and then present the corresponding decision rule. An upper bound on the performance of the resulting decision rule is also presented and utilized to select thresholds based on the desired probability of detection (PD) and probability of false alarm (PFA).

2. HYPOTHESIS TESTING OF TEMPORAL PROFILES

The initial problem is a 3-ary hypothesis testing scenario: the hypotheses are target plus noise, clutter plus noise, and background (here, clear sky) plus noise. (The fourth possibility, target plus clutter plus noise, will be treated here as simply clutter plus noise.) The background is modeled as the sum of a constant and temporally white background noise whose variance is constant for a given imaging system and can be experimentally measured. Very large or slow-moving clutter will appear as a slowly varying random signal plus the same noise process. In previous work [9] we showed that a simple temporal bandpass filter successfully eliminated both of these cases by using a threshold determined from the noise variance. Thus this filter can be used as a pre-processing stage. After pre-processing, the remaining pixels can reliably be considered to be either targets or potential false alarms caused by clutter from evolving cloud edges.

The time signature of a pixel affected by a point target will

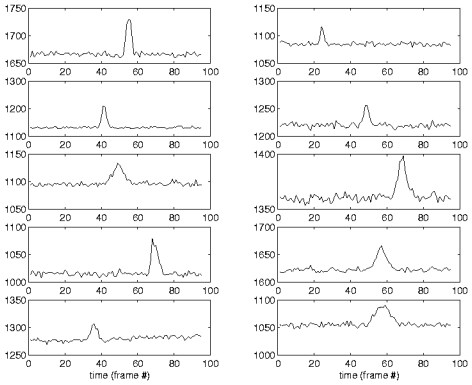

Figure 1: Temporal pixel profiles of real targets, extracted from infrared image sequences, are shown in left column. Simulated targets, created using the model with varied values of the unknown parameters, are shown in the right column.

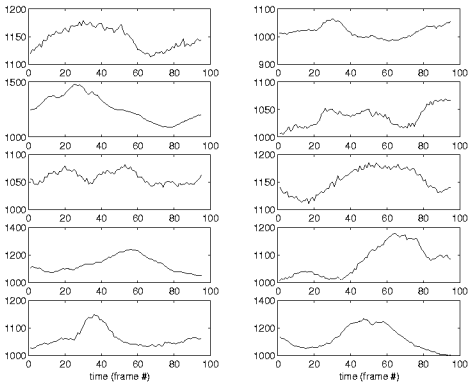

Figure 2: Plots of temporal profiles of challenging cloud clutter pixels.

have a pulse like shape. The width of the pulse will be inversely proportional to the speed of the target, whereas the intensity will be proportional to its strength. Assuming that the target is a true point source, the target profiles are actually dilated or contracted versions of one dimensional profiles of the point spread function (PSF) of the imager, which we have measured [10]. Thus we have a known deterministic model for the target signal with unknown parameters determining its velocity, intensity, background level, and time of arrival. We define *target intensity* as the highest deviation of any single signal sample from the background level over the entire signal. SNR is then defined as the ratio of target intensity to the background noise standard deviation. Figure 1 shows four target temporal profiles extracted from real image sequences in the left column and a set of target temporal profiles, simulated using the model, in the right column.

The other pixels which survive the pre-processing are generally caused by target-like clutter, from the edges of evolving clouds. Some typical difficult cloud clutter pixels are shown in Fig. 2. These pixels are expected to have temporal profiles with broader peaks than those seen in typical target profiles (Fig. 1). Through extensive analysis of our large database of real world IR sequences, we have determined that for cloud clutter these pixel profiles are well fit by a first order Markov model: a clutter pixel value at time k is modeled as the sum of the value of that pixel at time $k-1$ and a white Gaussian noise or error term. The magnitude of the standard deviation of the driving noise of the model describes the severity of the clutter in a scene. We define signal-to-clutter ratio (SCR) as the ratio of target intensity to the average clutter model standard deviation in the scene.

Based on these target and clutter models we developed a hypothesis test by constructing the following statistic for a generalized likelihood ratio test

$$\lambda(R) = \sum_{k=1}^{N} \left[\frac{[r(k) - r(k-1)]^2}{2\sigma_c^2} - \frac{[r(k) - f(k;\mathbf{p})]^2}{2\sigma_n^2} \right] \quad (1)$$

where $r(k)$ is the pixel temporal profile, $f(k;\mathbf{p})$ is the deterministic target model with unknown parameter vector $\mathbf{p}$, and σ_c^2 and σ_n^2 are the variances of the input noise to the Markov model and the background camera noise, respectively. The first term in the summation measures how well a given pixel temporal profile matches the Markov clutter model, while the second term measures how well it matches the target signature $f(k;\mathbf{p})$. The detection statistic is compared to a threshold to decide whether the pixel follows the target or the clutter model. The unknown parameters $\mathbf{p}$ of the target model must first be estimated from the data, for instance by a set of velocity-matched filters.

3. UPPER BOUND PERFORMANCE ANALYSIS OF GLRT

In this section we present a *perfect measurement* upper bound [11] on the performance of the GLRT presented in the previous section, by assuming knowledge or perfect estimates of the unknown parameters in $\mathbf{p}$, σ_n and σ_c. In practice, the unknown parameters would be estimated from the observed signal, causing a degradation in the performance of the test. To obtain receiver operator characteristic (ROC) curves for the test we need to determine the PDF of the detection statistic $\lambda(R)$, under the two hypotheses. A different method was required to analyze $\lambda(R)$ under each hypothesis.

3.1. PDF of $\lambda(R)$ under clutter present hypothesis

Under the assumption that the received or observed temporal profile $r(k)$ follows the clutter model, the detection statistic becomes

$$\lambda(R) = \sum_{k=1}^{N} \left[\frac{[w(k)]^2}{2\sigma_c^2} - \frac{[r(k) - f(k;\mathbf{p})]^2}{2\sigma_n^2} \right] \quad (2)$$

where $w(k)$ is the first temporal difference of $r(k)$. Writing $r(k)$ as $r(k) = r(0) + \sum_{j=1}^{k} w(j)$ gives

$$\sum_{k=1}^{N} \left[\frac{[w(k)]^2}{2\sigma_c^2} - \frac{[r(0) + \sum_{j=1}^{k} w(j) - f(k;\mathbf{p})]^2}{2\sigma_n^2} \right] \quad (3)$$

The only random quantities in the above expression are the $w(\cdot)$ terms. After considerable algebraic manipulation, $\lambda(R)$ can be approximated as a difference of quadratics of normal random variables. The PDF of $\lambda(R)$ can then be expressed as a weighted sum

of central Chi-Square distributions [12, 13] of the form

$$f(x) = \sum_{i=0}^{\infty} c_i \chi^2_{N+2i}(x/p) \tag{4}$$

where $\chi^2_k(x)$ is a central Chi-Square distribution with k degrees of freedom and the coefficients c_i and p can be determined from the values of σ_c and σ_n and the known signal $f(k;\mathbf{p})$. Details on this development are given in [14].

3.2. PDF of $\lambda(R)$ under target present hypothesis

Under the assumption that the received or observed signal $r(k)$ follows the target model, the detection statistic becomes

$$\lambda(R) = \sum_{k=1}^{N} \left[\frac{[r'(k)]^2}{2\sigma_c^2} - \frac{[n(k)]^2}{2\sigma_n^2} \right] \tag{5}$$

where $r'(k)$ is the first order temporal difference sequence given by

$$r'(k) = f(k;\mathbf{p}) - f(k-1;\mathbf{p}) + n(k) - n(k-1). \tag{6}$$

To obtain the PDF of the expression in Eq. 5 we used a characteristic function method. Here we make a slight change of notation, using n_k for $n(k)$ in Eq. 5. By viewing $\lambda(R)$ as a function of the N independent Gaussian random variables n_k, we express the characteristic function of $\lambda(n_1, n_2, \ldots, n_N)$ as:

$$\Phi_\lambda(\omega) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p_{n_1}(n_1) p_{n_2}(n_2) \ldots p_{n_N}(n_N) \exp[j\omega\lambda(n_1, n_2, \ldots, n_N)] dn_1 dn_2 \ldots dn_N$$

where $p_{n_k}(n_k)$ is the Gaussian distribution function of n_k. We can separate this N-dimensional integration into N 2-dimensional integrations since $\lambda(n_1, n_2, \ldots, n_N)$ only contains cross-terms between adjacent n_k. Thus $\Phi_\lambda(\omega)$ can be easily evaluated numerically. The PDF of $\lambda(R)$ can then be determined from the magnitude of the inverse Fourier transform of the characteristic function $\Phi_\lambda(\omega)$.

3.3. ROC Curves

Using the PDF's of $\lambda(R)$ derived in the previous section, we computed ROC curves that can be used to gain insight into the performance of the test under different situations and to select thresholds depending on the desired PD and PFA. In Fig. 3 we plot PD vs PFA for different velocity targets of the same intensity. The background noise and clutter model standard deviations (σ_n and σ_c) were set to 3 and 4.5 respectively, which are typical values for our infrared sequences. The target intensity parameter was set to 12 (SNR = 4, SCR = 2.67), which is the lower SNR limit of the targets we are attempting to detect. Note that slower targets are more difficult to detect than faster ones, since they have smaller temporal differences and essentially "look more like" clutter.

4. EXAMPLE RESULTS

In this section, we demonstrate the effectiveness of the proposed algorithm by testing it on real infrared image sequences containing airplanes flying across the scene at long range. The sequences

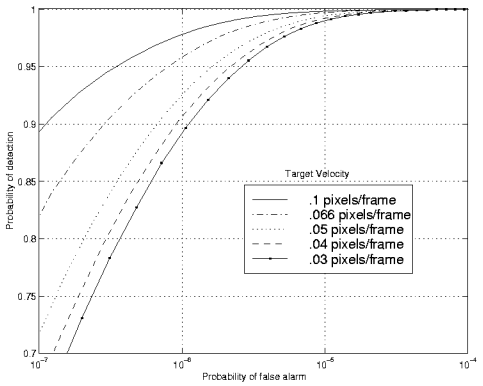

Figure 3: ROC curves for targets of varying velocities, at fixed SNR = 4 and SCR = 2.67.

Figure 4: Sample image from infrared image sequence. The outlined areas designate the location of the targets.

were acquired using a PtSi infrared camera with a 320×244 pixel focal plane array, imaging at 30 frames per second. A single image from a sample sequence is shown in Fig. 4. First all pixels except targets and potential clutter false alarms were eliminated through the pre-processing stage. For this sequence all but 105 pixels (.13%) were eliminated. A binary image indicating the location of the pixels that passed the pre-processing step is shown in Fig. 5. After applying the GLRT to these pixels and thresholding, only 7 pixels were above the threshold. Using the ROC curve shown in Fig. 3, we selected a threshold to ensure a worst case PFA of 10^{-5} for targets with velocities of .03 pixels/frame or higher at an SNR of 4. This corresponds to a worst case PD of approximately 99%. A binary image showing the location of the pixels that were above the threshold is shown in Fig. 6. Notice that both targets are successfully detected with zero false alarms.

The algorithm was applied to 12 image sequences, containing a total of 17 targets. Using the same threshold (selected as described above) for all sequences, 13 of the 17 targets were successfully detected with a total of 4 false alarms. These false alarms were caused by moving clouds that are very thin spatially, resulting in temporal profiles that look very similar to those of target pixels. These few situations could then benefit from further spatial processing, discussed in the following section.

Figure 5: Binary image indicating in white the location of the pixels that passed the pre-processing step.

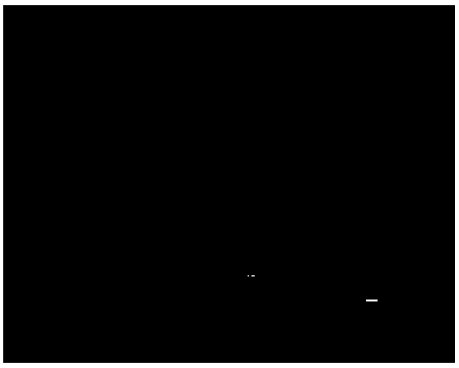

Figure 6: Binary image indicating in white the location of the pixels that were above the threshold after applying the GLRT. Notice that both targets are successfully detected with zero false alarms.

5. SPATIAL HYPOTHESIS TEST

In severe clutter situations, additional spatial processing can be utilized to eliminate false alarms. A spatial hypothesis test, similar in nature to the temporal one described in Sec. 2, can be developed using spatial models for the target and cloud clutter. Assuming that the targets are true point targets, the target spatial profiles can be modeled using the 2-D PSF of the imager. Because the clouds are moving slowly across the scene, a 1-D spatial profile is essentially a subsampled version of the temporal profile. Therefore an appropriate Markov-type model will fit the spatial signature of cloud clutter, over small spatial regions. Using these models, we developed a spatial hypothesis test, which is only applied to those pixels that were above the threshold after the temporal GLRT. We have obtained preliminary results but space limitation preclude including them here. Examples demonstrating that this additional spatial processing can eliminate false alarms in severe clutter situations and increase the probability of detection for weak targets will be included in the oral paper presentation.

6. CONCLUSIONS

In this paper we present an algorithm for the detection of small moving objects in image sequences that also contain moving and evolving clutter. Using temporal models for the targets and clutter on a single pixel basis, we developed a hypothesis testing procedure and derived and analyzed the performance of the corresponding decision rule. We demonstrated the effectiveness of the resulting algorithm on real infrared image sequences containing airplanes flying at long range. Future work will concentrate on completion of the spatial hypothesis test introduced in the previous section, as well as application of the approach to other scenarios such as biological or biomedical imaging.

7. REFERENCES

[1] D. S. K. Chan, D. A. Langan, and D. A. Staver, "Spatial processing techniques for the detection of small targets in IR clutter," *Proc. SPIE*, vol. 1305, pp. 53–62, 1990.

[2] C. D. Wang, "Adaptive spatial/temporal/spectral filters for background clutter suppression and target detection," *Optical Engineering*, vol. 21, no. 6, pp. 1033–1038, 1982.

[3] I. Reed, R. M. Gagliardi, and H. M. Shao, "Application of three-dimensional filtering to moving target detection," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-19, no. 6, pp. 898–905, 1982.

[4] D. Braunweiter and N. Banh, "Detection of moving subpixel targets in infrared clutter with space time filtering," *Proc. SPIE*, vol. 1481, pp. 73–83, 1991.

[5] Y. Barniv, "Dynamic programming solution for detecting dim moving targets," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-21, no. 6, pp. 144–156, 1985.

[6] J. Arnold and J. Pasternack, "Detection and tracking of low-observable targets through dynamic programming," *Proc. SPIE*, vol. 1305, pp. 206–217, 1990.

[7] S. D. Blostein and T. S. Huang, "Detecting small, moving objects in image sequences using sequential hypothesis testing," *IEEE Transactions on Signal Processing*, vol. SP-19, no. 7, pp. 1611–1629, 1991.

[8] J. Silverman, J. M. Mooney, and C. E. Caefer, "Temporal filters for detecting weak slow point targets in evolving cloud clutter," *IR Physics and Technology*, vol. 37, pp. 695–710, 1996.

[9] A. P. Tzannes and D. H. Brooks, "Temporal filters for point target detection in IR imagery," *Proc. SPIE*, vol. 3061, pp. 508–520, 1997.

[10] A. P. Tzannes and J. M. Mooney, "Measurement of the modulation transfer function of infrared cameras," *Optical Engineering*, vol. 34, no. 6, pp. 1808–1817, 1995.

[11] H. L. Van Trees, *Detection, Estimation and Modulation Theory*. Wiley, 1968.

[12] H. Ruben, "Probability content of regions under spherical normal distributions, IV: the distribution of homogeneous and non-homogeneous quadratic functions of normal variables," *Annals of Mathematical Statistics*, vol. 33, pp. 542–570, 1962.

[13] N. L. Johnson and S. Kotz, *Continuous Univariate Distributions - 2*. Boston: Houghton Mifflin Company, 1970.

[14] A. P. Tzannes and D. H. Brooks, "Detection of small moving objects in image sequences by temporal hypothesis testing." submitted for publication to *The Journal of the Optical Society of America A*, August 1998.