Wavelet Based Detection and Modified Pipeline Algorithm for Multiple Point Targets Tracking in InfraRed image sequences

Mukesh A. Zaveri Anant Malewar[‡] S. N. Merchant Uday B. Desai
SPANN Lab, Electrical Engineering Dept., IIT Bombay - 400076. [*]

Abstract

A wavelet based technique for multiple point targets detection in InfraRed image sequences in presence of clutter is proposed. Most existing approaches assume target of several pixels size or of Gaussian shape with variance of 1.5 pixel. We develop detection and tracking algorithm for single pixel targets. We also propose a modified pipeline algorithm for tracking multiple single pixel targets with large motion (± 20 pixels) in an Infrared image with occlusion due to IR clouds and background noise.

1 Introduction

Most of IRST (Infrared Search and Track) systems assume Gaussian distributed size or blob size target [21]. In such cases, matched filtering ([11], [13]) approach can be used for target detection. These approaches cannot be extended to single pixel targets (i.e. point targets). Another problem with point targets is that intensity alone cannot be used as matching criterion for detection because the target intensity varies continuously with changing distance between the imaging device and target. Moreover, clutter (clouds) also causes change in intensities. In such a situation, detection of point target requires integration of target intensity over multiple frames ([10], [18]) and exploitation of motion.

Adaptive algorithm based on TDLMS (Two Dimensional Least Mean Square) is proposed for target detection in [2]. In our simulations, we found that TDLMS algorithm is very slow in case of large frame size. Recently, nonlinear filters have been widely used for clutter suppression ([15], [16]). Sequential probability ratio test [5] and dynamic programming (DPA) [4] based approaches proposed by Blostein and Barniv are computationally expensive. In [8] different models for target and clutter are specified and a matched filter is used to detect the target in the presence of clutter. Different methods based on differencing two images are presented in [7] to suppress the background and retain only the target. Temporal based algorithm using triple temporal filter (TTF) for simultaneous detection and tracking of a point target in consecutive IR frames based on six parameters is presented in [6]. For naval surveillance, a method is proposed in [17] to remove the temporally correlated background clutter based on discrete KLT transform of the vectors representing the columns of the image. Multiple blob size targets detection was carried by exploiting temporal decomposition in [12].

In our proposed method we apply temporal decomposition to detect multiple targets in the presence of clutter using change detection map. Output of this stage is given to the tracking phase as candidate target list.

Along with target detection, tracking of multiple targets in the presence of clutter is paramount in any IRST system. Different methods to track multiple targets based on multiple hypotheses testing (MHT) have been proposed in [19] and [20]. Target tracking based on probabilistic data association (PDA) and Joint PDA is described in [3], [14] and [22]. All these methods are computationally expensive and have little scope for real time application.

A pipeline algorithm proposed by Wang et.al. [9] is able to detect and track pixel-sized moving target in noisy environment having continuous and smooth trajectories. The disadvantage with this approach is that it works with only slow target movement, typically 1-2 pixels per frame. Moreover, computationally it is expensive for large image frame, since it forms a pipe for every pixel in the frame. Another problem with this method is that it is not able to track a target in case of occlusion caused by clouds over several frames. We propose a major modification of [9] whereby we are able to overcome the above drawbacks. In the proposed method, constant acceleration Kalman filter is used to predict the position in next frame. A search window is formed around this predicted position. A pipe is formed at all pixels only inside the search window and not the whole frame. As will be seen later, this helps in handling large target motion (± 20 pixels), occlusion due to clouds, and binary noise.

Our algorithm for point target detection and tracking is as follows: the first phase detects targets in the image sequence using proposed wavelet based technique and in the second phase multiple targets are tracked using modified pipeline algorithm, wherein a constant acceleration Kalman

[*] ‡Mr. Malewar is now working with PACE Software, Pune. Emails: mazaveri@ee.iitb.ac.in, anantmalewar@yahoo.com, merchant@ee.iitb.ac.in, ubdesai@ee.iitb.ac.in

filter is used for tracking. This tracking algorithm updates or deletes existing tracks and may initiate new tracks using the output of the detection phase. The paper is organized as follows. In Section 2, the wavelet based target detection scheme is introduced. The modified pipeline algorithm is described in Section 3. Simulation results are presented in Section 4.

2 Wavelet Based Detection

A target appears as a point in an image sequence when it is very far from the sensor. The task now is to detect multiple point targets and track them in the presence of occlusion and noise. In such cases segmentation based techniques will fail due to point nature of the targets. It is not possible to classify point target from noise pixel. To detect target in such an environment we exploit a change detection algorithm based on temporal multiscale decomposition.

2.1 Temporal Multiscale Decomposition

A temporal multiscale decomposition allows one to detect and to characterize various dynamical behavior of the elements present in a scene. To detect small moving objects in a clutter background, a longer temporal integration is required in absence of any texture information. All these temporal signals can be considered as constant signals, i.e. there is no frequency content in the absence of any intensity variation or any moving object. Wavelet transform is widely used to give better time-frequency localization. A wavelet basis is composed of a family of functions adjusted by two parameters: one for the position (in time) b, the other for the scale, a. The wavelet basis $\emptyset_{mn}(t)$ can be written as follows:

$$\emptyset_{mn}(t) = a^{-m/2}(a^{-m}t - nb) \tag{1}$$

The original temporal signal is denoted C^0. At a given level k the signal C^k, called approximation signal, is split up into two terms : a new approximation signal at a coarser scale C^{k+1}

$$C^{k+1}(i) = \sum_n H(n-2i)C^k(n) \tag{2}$$

(H is equivalent to a low pass filter) and a signal coding the difference in the information, D^{k+1}

$$D^{k+1}(i) = \sum_n G(n-2i)C^k(n) \tag{3}$$

(G is equivalent to high pass filter). D^1 characterizes high temporal frequencies components. Following levels D^2, $D^3 \ldots$ correspond to lower frequency bands. The Harr basis is used in experimentation, since, as the number of filter coefficients increases, large number of image frames are required for temporal multiscale decomposition, which introduces further delay in making decision. The advantage of temporal multiscale decomposition is that no preprocessing technique, like spatial smoothing (low pass filter or median filter) is required. Preprocessing has the disadvantage that single pixel targets may get eliminated. The temporal multiscale decomposition allows to build temporal intensity change maps at various temporal scales. These maps indicate whether there is temporal change or not. These binary maps are intermediate decision maps representing presence or absence of temporal changes at each resolution level.

A two-hypotheses likelihood ratio test is applied to validate temporal changes at each scale. By inclusion of hypotheses and, hence, likelihood ratio test, the motion detection issue is solved in statistical frame work. Two competing hypotheses are compared : hypothesis H_0 (no temporal change at p) and hypothesis H_1 (temporal change at p). The log-likelihood ratio corresponding to hypotheses H_1 and H_0 is derived and the decision step is formalized as :

$$\psi^k(p) \underset{H_1}{\overset{H_0}{\substack{<\\ \geq}}} \lambda \tag{4}$$

where $\psi^k(p)$ is the resulting expression of the log-likelihood ratio in the maximum likelihood sense. λ is a threshold which may be inferred from tables of statistical laws. At each scale k,

$$\psi^k(p) = \frac{1}{2\sigma_k^2}\left[\frac{1}{N}\left(\sum_{i=1}^{N} D^k(p_i)\right)^2 + \frac{1}{\sum x_i^2}\left(\sum_{i=1}^{N} x_i D^k(p_i)\right)^2 + \frac{1}{\sum y_i^2}\left(\sum_{i=1}^{N} y_i D^k(p_i)\right)^2\right] \tag{5}$$

$\psi^k(s)$ follows χ^2 distribution with three degrees of freedom. N is the size of the window in terms of pixels centred at point p_i and (x_i, y_i) indicates the relative location of pixels w.r.t. the centre of window. σ_k^2 is variance of the pixel intensity within the window. It is important to note that the above likelihood ratio test is applied to difference in information image, D^k, and not to the original image. At least three levels of the wavelet transform are required to allow us to correctly discriminate the various dynamical behavior present in the scene. The following heuristics are used to characterize three specific dynamical behavior.

- If a pixel at time t is detected as temporal change for at least three successive temporal scales, then with very high probability it is a moving object.
- If a pixel at time t is never detected as temporal change at any temporal scale, then it is static.

Figure 1: Change Detection Map at frame number 2 (IR clips:1)

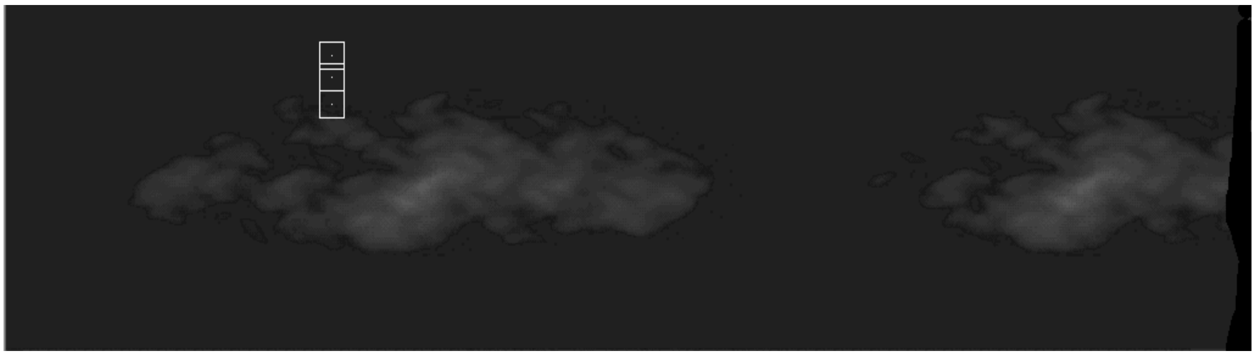

Figure 2: Target Identification after Postprocessing of Change Detection Map at frame number 2 (IR clips:1)

- If a pixel at time t is detected as temporal change in at most two successive temporal scales, it is likely to be a temporary temporal change, and most likely due to noise.

2.2 Postprocessing for Target Detection

In order to make the detection scheme robust to clutter and noise, a post processing technique is proposed. Moreover, for point target detection, postprocessing will reduce false alarm and misdetection. Change detection map is segmented and then all segments having size larger than a threshold defined by δ_{th} are removed. From a segmented image, candidate target list is prepared and used for further processing. List contains information about size and centroid location of each segment. Special attention is required in three different cases which arises in IR image sequences:

- The clouds may be moving with a significant speed, i.e. background is not stationary and continuously varying.
- Clouds are scattered and appear like blob sized targets.
- Edges of any undesired object or clouds in image sequences significantly contribute to change detection map.

The edge effects and small size clutter which appear like a small target are eliminated using the following procedure:

1. *Local contrast* $lc(x_n, y_n)$: let us consider an image pixel (x_n, y_n) which is candidate point target from a list prepared after segmentation, belonging to a segment R_i of the segmented image. $lc(x_n, y_n)$ is defined as
$$lc(x_n, y_n) = \mid I(x_n, y_n) - \frac{1}{s_i} \sum_{(x_m, y_m)\ \epsilon\ A_f} I(x_m, y_m) \mid \tag{6}$$
where $I(x_n, y_n)$ is the gray level value of the pixel at (x_n, y_n), s_i is the size of neighborhood window in terms of pixels and A_f is defined as
$$A_f = \{(x_j, y_j) \mid (x_j, y_j)\ \epsilon\ N_r(x_n, y_n) \mid (x_n, y_n) \neq (x_j, y_j)\}$$
where N_r is the neighborhood window defined by a circle of radius r centered at (x_n, y_n). $lc(x_n, y_n)$ is compared with predefined threshold ρ. If it crosses the threshold it may be a point target. It ensures that blob sized scattered cloud or edge effects will be removed.
2. If the above threshold is crossed at point (x_n, y_n) then to avoid the problem of small size clutter, its intensity is compared with intensity of pixels within eight-connected region only. Consider a pixel at (x_n, y_n),

then we accept or reject it as a candidate target as per the following: if $| I(x_n, y_n) - I(x_m, y_m) | < \varepsilon$ for $\forall(x_m, y_m) \in N_8(x_n, y_n), m \neq n$, then reject. Here $N_8(x_n, y_n)$ is the eight-connected neighbor of (x_n, y_n).

3. Output of the above step gives isolated point target. Temporal decomposition and likelihood ratio test ensure that this isolated point is not due to noise.

Once moving targets are detected by the above method, tracking module updates existing tracks and a new temporal filter is created if none of the existing tracking filter can be associated with this moving targets. If no data is available during ten iterations, i.e. ten successive predictions without any correction, the temporal filter is killed.

3 Tracking using Modified Pipeline Algorithm

The pipeline algorithm proposed in [9] is able to detect and track targets having 1-2 pixel movement per frame. To track a target with large movement (± 20 pixels per frame), the window size needs to be increased and hence it leads to

- False alarm due to increased search neighborhood in continuity filtering algorithm (CFA).
- Consequently, taking centroid position of Temporal Window Column (TWC) as a detected target position will not be correct.
- Number of additions per pixel increase by $O(N^2)$ with increase in maximum detectable target movement, where N is the number of pixels in a frame.

To overcome these, we propose a modified pipeline algorithm. We propose the use of a variable slope pipe as shown in Figure 3. Pipeline algorithm consists of two major components: (i) fixed length TEST PIPE of temporally adjacent image frames. At each iteration a frame from the bottom of the test pipe is discarded and new frame is added to top of the pipe and other frames are shifted in pipe by one position. (ii) AND PIPE (AP), consists of two image frames and a single blank frame called Target Frame (TF). The CFA uses the AND PIPE to test continuity of the target pixel. The algorithm is as described below.

1. Initialize the AND PIPE (AP) by adding a frame to the top at each cycle.
2. Initialize the TEST PIPE (TP) with $n-1$ frame cycles, where n is the number of frames in TP.
3. At each time step,
   (a) A constant acceleration based Kalman filter is used in a predictor mode, which predicts the target position in the next frame.
   (b) i. Update the AP.
       ii. Apply the CFA in AP.
   (c) i. Update the TP by adding the output frame of the AP in each cycle to the top of the TP.
       ii. Form a search window around the predicted position given by the Kalman filter. Apply the TWC at each pixel (X_i, Y_i) in the search window and sum the intensities. This greatly reduces the computational load. This sum is given by:

$$S(X_i, Y_i) = \sum_{k=1}^{n} \sum_{x=-w/2}^{w/2} \sum_{y=-l/2}^{l/2} I(X_i + x, Y_i + y, k) \tag{7}$$

       where n, w and l are the dimension of the TWC in the TP.
       iii. If sum is greater than threshold (determined by possible target intensities) then go to Step 3(d), else consider it as occlusion and go back to Step 3(a) for the next time step.
   (d) Compute the intensity centroid. Find the Euclidean distance between the current position and the previous position and compare it against a threshold. If it is below the threshold then, accept as trajectory pixel and record in the TF else reject.

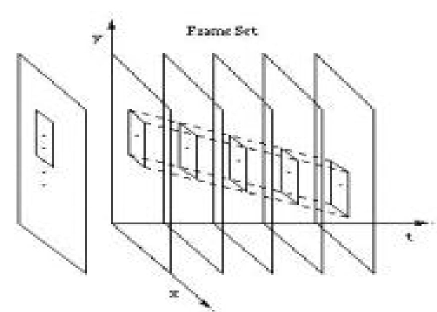

Figure 3: Modified Test Pipe

The modified pipeline algorithm requires coarse estimate of initial velocity. To overcome this, the output of target detection phase is used in following way.

- The candidate target list is formed using the first two frames.
- Using nearest neighbor technique with maximum velocity constraint, associate candidate from one list to another list, which provides estimation of velocity.

- Initiate a new target track if association is not possible for any candidate target in the list.

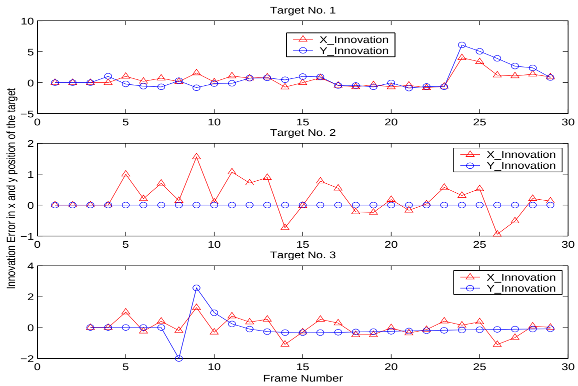

Figure 6: Innovation Error

4 Simulation and Results

Synthetic InfraRed images are generated using real time temperature data [1]. Intensity at different points in images is function of temperature, surface property and other environmental factors. We are using Gardner's method to synthesize InfraRed clouds. For simulation purpose, the generated frame size is 1024×256 and the target movement is ± 20 pixels per frame. In simulation, eight frames are used to detect the target, i.e., the target is seen for a short duration for detection. No assumptions are made on the velocity of the target.

Figure 1 represents temporal change at each pixel, found using the wavelet based technique. It also shows clutter. The clutter is removed by postprocessing of the change detection map. The detected candidate targets are shown in Figure 2. We have also tested the proposed target detection algorithm on real IR image sequence. The results are not included in this paper due to space constraint.

Figure 4 shows the trajectories of targets formed using the modified pipeline algorithm. We also tested our proposed modified pipeline algorithm with IR images having binary noise. The proposed algorithm performs well even in the presence of binary noise as shown in Figure 5. Use of Kalman filter helps us to take care of occlusion of target over number of frames (see frame no. 17 to frame no. 23 in Figures 4 and 6 for target no. 1). The innovation error in x and y positions of the targets is illustrated in Figure 6. We define the signal to clutter + noise ratio (SCNR) as

$$SCNR = 10 \log \frac{(S_t - m_0)^2}{\sigma_0^2} \tag{8}$$

where S_t is the minimum intensity value at a pixel in the presence of target, m_0 is the average value of the clutter plus noise and σ_0^2 is clutter plus noise power. With a proper choice of ε and δ_{th} the proposed scheme works very well with SCNR $\geq$ 2 dB.

5 Conclusions

For large frame size and large target movement, wavelet based detection scheme performs well in the presence of clutter and noise. The performance of proposed detection and tracking technique is satisfactory even in the presence of multiple targets, clutter and binary noise.

Acknowledgments

The authors would like to acknowledge Avinash Pandit and Shanker More for providing the IR video clips. This research project is supported by Department of Extramural Research and Intellectual Property Rights, DRDO, New Delhi.

References

[1] S. M. A. Pandit, S. More and U. B. Desai. IR scene simulation, search and tracking system. *Technical report, SPANN Lab, Electrical Engg. Dept., IIT, Bombay - 400076*, Jan. 2002.

[2] S. K. A.S. Elfishawy and A. Abutaleb. Adaptive algorithms for change detection in image sequence. *IEEE Signal Processing*, 23(2):179–191, May 1991.

[3] Y. Bar-shalom and T. E. Fortmann. Tracking and data association. *Academic Press*, 1989.

[4] Y. Barniv. Dynamic programming solution for detecting dim moving targets. *IEEE Transactions on Aerospace and Electronic Systems*, 21(1):144–156, Jan. 1985.

[5] S. Blostein and T. Huang. Detecting small, moving objects in image sequences using sequential hypothesis testing. *IEEE Transactions on Signal Processing*, 39(7):1611–1629, 1991.

[6] J. S. Charlene E. Caefer and J. M. Mooney. Optimization of point target tracking filters. *IEEE Transactions on Aerospace and Electronic Systems*, 36(1):15–25, Jan. 2000.

[7] B. V. K. D. Casasent and Y. Lin. Subpixel target detection and tracking. *SPIE Intelligent Robots and Computer Vision: Fifth in a Series*, 726:206–220, 1986.

[8] D. A. L. David S.K. Chan and D. A. Staver. Spatial processing techniques for the detection of small targets in ir clutter. *SPIE Signal and Data Processing of Small Targets*, 1305:53–62, 1990.

[9] E. S. M. Gan Wang, Rafel M. Inigo. A pipeline algorithm for detection and tracking of pixel-sized target trajectories. *SPIE Signal and Data Processing of Small Targets*, 1305:167–177, 1990.

[10] H. Im and T. Kim. Optimization of multiframe target detection schemes. *IEEE Transactions on Aerospace and Electronic Systems*, 35(1):176–187, Jan. 1999.

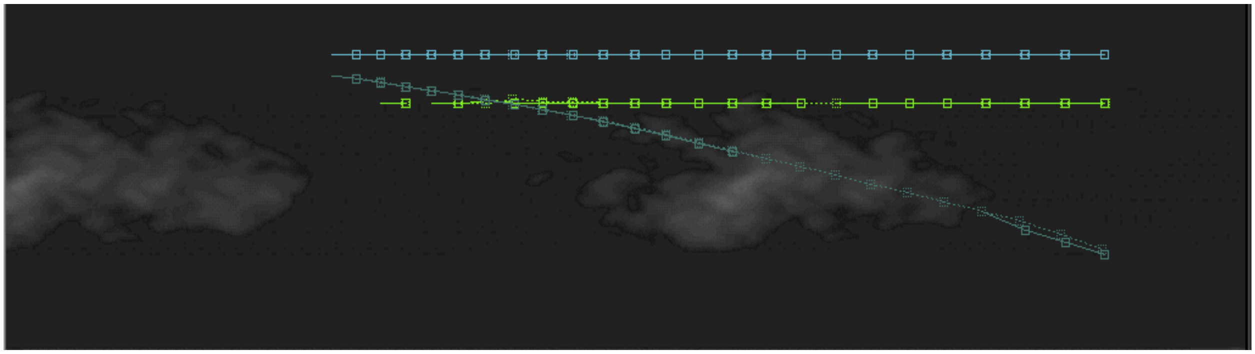

Figure 4: Target Trajectories at frame number 26 (IR clips:1)

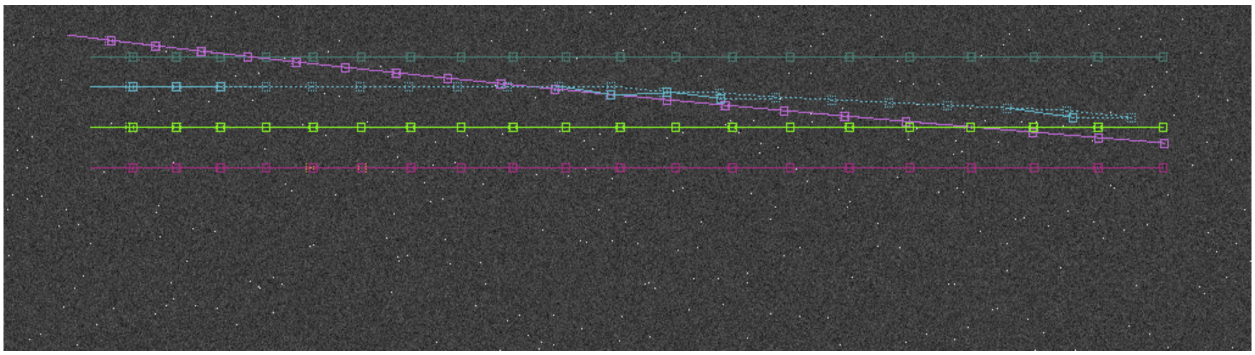

Figure 5: Example of Pipeline Algorithm with Binary Noise (IR clips:2)

[11] R. M. G. Irving S. Reed and L. B. Stotts. A recursive moving-target-indication algorithm for optical image image sequences. *IEEE Transactions on Aerospace and Electronic Systems*, 26(3):434–440, May 1990.

[12] P. B. . V. R. Jean-Michel LÉTANG. Robust motion detection with temporal decomposition and statistical regularization. *INRIA Report*, Oct. 1995.

[13] G. A. Lampropoulos and J. F. Boulter. Filtering of moving targets using sbir sequential frames. *IEEE Transactions on Aerospace and Electronic Systems*, 31(4):1255–1267, Oct. 1995.

[14] M. Lee and Y. Kim. New data association method for automotive radar tracking. *IEE Proceedings - Radar, Sonar Navigation*, 148(5):297–301, Oct. 2001.

[15] H. Leung. Nonlinear clutter cancellation and detection using a memory-based predictor. *IEEE Transactions on Aerospace and Electronic Systems*, 32(4):1249–1256, Oct. 1996.

[16] H. Leung and A. Young. Small target detection in clutter using recursive nonlinear prediction. *IEEE Transactions on Aerospace and Electronic Systems*, 36(2):713–718, April 2000.

[17] A. B. M. Diani and G. Corsini. Novel background removal algorithm for navy infrared search and track systems. *Optical Engineering*, 40(8):1729–1734, Aug. 2001.

[18] J. Z. Paul Wei and W. Ku. Analysis of multiframe target detection using pixel statistics. *IEEE Transactions on Aerospace and Electronic Systems*, 31(1):238–247, Jan. 1995.

[19] D. B. Reid. An algorithm for tracking multiple targets. *IEEE Transactions on Automatic Control*, 24(6):843–854, Dec. 1979.

[20] R. D. S. Blackman and T. Broida. Multiple hypothesis track confirmation for infrared surveillance systems. *IEEE Transactions on Aerospace and Electronic Systems*, 29(3):810–824, July 1993.

[21] A. R. Vishal Markandy and S. Wang. Motion estimation for moving target detection. *IEEE Transactions on Aerospace and Electronic Systems*, 32(3):866–874, July 1996.

[22] B. Zhou and N. Bose. Multitarget tracking in clutter: Fast algorithms for data association. *IEEE Transactions on Aerospace and Electronic Systems*, 29(2):352–363, April 1993.