AIR-BORNE APPROACHING TARGET DETECTION AND TRACKING IN INFRARED IMAGE SEQUENCE

Mukesh A. Zaveri   S. N. Merchant   Uday B. Desai

SPANN Lab, Electrical Engineering Dept., IIT Bombay - 400076.
Email address : [mazaveri,merchant,ubdesai]@ee.iitb.ac.in

ABSTRACT

Detection and tracking of approaching targets in an infrared (IR) image sequence is important for surveillance applications. In this paper an algorithm is proposed which provides a complete solution (track while scan) for detection and tracking for IRST system. The proposed method uses only motion as a cue to detect the target. Detection is followed by tracking. In a real scenario, the movement of a target is arbitrary and no apriori information is available. We propose a tracking method which tracks maneuvering and non-maneuvering targets simultaneously using a filter bank. The switch-over amongst the filters is based on a single-step decision logic.

1. INTRODUCTION

In literature, various approaches have been proposed for detection and tracking blob sized or Gaussian distributed targets. Most of these use match filtering [1]. On the other hand, very few results are available for detection and tracking of approaching targets. Detection and tracking of approaching target is a very challenging task because the intensity of the target and that of the cloud varies with time. Even the shape and size of the cloud and target may vary from one time instant to another time instant. Because of these reasons, the methods based on spatial processing, [2, 3] and optical flow [4] do not help. Methods based on background registration [5] or background extraction [6] may not be feasible due to the non-stationary nature of the background. Detection of approaching targets based on segmentation of each frame [7, 8] will not be viable for low contrast image sequence and lack of any texture information.

In this paper we propose a method to detect and track approaching targets in IR image sequences using motion as the only cue. This work is based on our earlier work on point target detection and tracking [9, 10]. Our earlier approach [9] performs very well for multiple point target detection in the presence of clutter (evolving clouds). Moreover, it does not require any preprocessing to detect target. In the present work, we suppress the background and this helps to improve the Signal to (Clutter+Noise) Ratio (SCNR) in the sequence. We have used wavelet transform based temporal filter for motion detection. The main thrust of the proposed approach is to detect the approaching targets under low contrast, without using any texture information or any feature other than motion. After detection, we track the targets using a filter bank approach proposed by us in [10]. The key feature of the filter bank approach is its ability to track maneuvering as well as non-maneuvering trajectories in the absence of any apriori information about the target dynamics.

2. APPROACHING TARGET DETECTION

2.1. Background Suppression

In air-borne video sequences targets appear to be very small compared to the background. Here, background consists of cloud and sky. As images are captured from a moving platform, the background has dominant motion compared to the target. Thus, background suppression significantly helps in improving SCNR. Due to the nonstationary (continuously varying) nature of the background, existing methods [6] do not yield acceptable results. We have developed a max-median filter [11] for this task. To obtain max-median image, a window of size $n \times n$ is run over an image. Median is found for column vector and row vector and two diagonal vectors, with respect to center pixel of the window. Next, the intensity value for the center pixel of the window is replaced by a maximum of these median values found.

For max-median filters, with 3×3 window, $I(x, y)$ is replaced by a value

$$\tilde{I}(x, y) = \max(med_1, med_2, med_3, med_4) \qquad (1)$$

where

$$med_1 = med(I(x, y-1), I(x, y), I(x, y+1))$$
$$med_2 = med(I(x-1, y), I(x, y), I(x+1, y))$$
$$med_3 = med(I(x-1, y-1), I(x, y), I(x+1, y+1))$$
$$med_4 = med(I(x+1, y-1), I(x, y), I(x-1, y+1))$$

The absolute difference image is obtained using the original image frame and the filtered output frame from the max-median filter. Now, a temporal filter is applied to the sequence of these difference image frames instead of original image frames.

2.2. Temporal Filter

We use wavelet transform based temporal filter for detecting moving targets. The wavelet transform enables one to detect and characterize the dynamical behavior of the elements present in the scene. We have used the Harr wavelet to reduce delay in detection. The temporal filtering facilitates the construction of the intensity change maps which indicate whether there is temporal change or not. A two-hypotheses likelihood ratio test is then applied to validate the temporal changes at each scale. By exploiting the likelihood ratio test, motion detection issue is solved in a statistical frame work. This hypotheses testing will give a temporal change detection map (binary map) which is further processed to reduce false alarm. Two hypotheses are defined, namely, H_1 corresponding to 'temporal change' at (x, y) and H_0 corresponding to 'no

0-7803-8554-3/04/$20.00 ©2004 IEEE.

temporal change'. The hypotheses are applied to difference in information signal D. Window of size of $n \times n$ is considered around the central pixel (x, y), where temporal change test is applied. N is the total number of pixels in $n \times n$ window. For simulation, the value of n is chosen to be 3. The log-likelihood corresponding to hypotheses H_1 and H_0 is derived and the decision step is formalized [12]:

$$\psi^k(p) \underset{H_1}{\overset{H_0}{\lesseqgtr}} \lambda \tag{2}$$

where $\psi^k(p)$ is the resulting expression of the log-likelihood ratio in the maximum likelihood sense at scale k,

$$\psi^k(p) = \frac{1}{2\sigma_k^2}\left[\frac{1}{N}\left(\sum_{i=1}^{N} D^k(p_i)\right)^2 + \frac{1}{\sum x_i^2}\left(\sum_{i=1}^{N} x_i D^k(p_i)\right)^2 + \frac{1}{\sum y_i^2}\left(\sum_{i=1}^{N} y_i D^k(p_i)\right)^2\right] \tag{3}$$

$\psi^k(p)$ follows a χ^2 distribution with three degrees of freedom. λ is a threshold which may be inferred from tables of statistical laws, (x_i, y_i) indicates the relative location of the pixel with respect to the center of the window. σ_k^2 is the variance of the pixel intensity within the window.

2.3. Clutter Removal

In order to make the detection scheme robust to clutter and noise, post processing is incorporated. In [9] the post processing step is used to detect only the point target. This step needs to be modified for approaching target detection. To merge nearby small regions in the change detection map, binary morphological operation, namely, closing (dilation followed by erosion) is performed on the change detection map. Next, the change detection map is segmented and all segments having a size larger than a threshold defined by δ_{th}, are removed. The edge effects and small size clutter which appear like small targets are eliminated by comparing local contrast $lc(x_n, y_n)$ with a threshold ρ. If it crosses the threshold it will be a moving target. Here, $lc(x_n, y_n)$ is defined as

$$lc(x_n, y_n) = \mid I(x_n, y_n) - \frac{1}{s_i} \sum_{(x_m, y_m)\, \epsilon\, N_w} I(x_m, y_m) \mid \tag{4}$$

where $I(x_n, y_n)$ is the gray level value of the pixel at (x_n, y_n) in the original image, s_i is the number of pixels in the neighborhood window centered at (x_n, y_n). N_w represents the neighborhood window.

3. APPROACHING TARGET TRACKING

The detection module provides the list of detected targets to the tracking module. This list is used as an available observation set for a given time instant. We use our earlier proposed method to track the approaching targets using multiple filter bank [10]. The filter bank consists of different types of filters. For example, in a bank of two filters, one could be a constant velocity filter for tracking non-maneuvering targets and the other could be based on a maneuver model for tracking maneuvering targets. We have used constant acceleration (CA) and Singer's maneuver model (SMM) for our simulations.

In the filter bank all the filters run in parallel. Nevertheless, at any given time instant, only one filter is designated as an active filter - this is the filter whose output is selected as the predicted position of the target. The decision on which filter becomes the active filter is based on minimum average innovation (the explicit procedure is described a little later in this section). To update existing tracks and to initiate new track, we need to select candidate observations and an algorithm to assign an observation to a track. The candidate observations are selected as those that fall within the validation gate (window) formed around the predicted position obtained from the active filter. Next, Munkres' optimal data association algorithm [13] is used for assigning a candidate observation to a track. At every time instant, an observation that is assigned by Munkres' algorithm, is given to all the filters in the filter bank. The filters now update their states independent of each other.

Selection of active filter: Let $\mathbf{z}(m)$ be an observation at time m assigned to a particular track by the data association algorithm. The innovations is defined by $\tilde{\mathbf{z}}(m) = \mathbf{z}(m) - \hat{\mathbf{z}}(m|m-1)$, where $\hat{\mathbf{z}}(m|m-1)$ is the predicted measurement. The innovation based measure is defined as

$$v(m) = \{\tilde{\mathbf{z}}^T(m)\mathbf{S}^{-1}(m)\tilde{\mathbf{z}}(m)\} \tag{5}$$

where S (a diagonal matrix) is the innovation covariance matrix. Next, an averaged accumulated innovation based measure is computed as

$$e_i(k) = \frac{1}{s}\sum_{m=k-s+1}^{k} v_i(m) \tag{6}$$

where $v_i(m)$ is the innovations based measure obtained using (5) for the i-th filter i and time instant m and s is size of the sliding window. Note that (6) is computed for all filters i, and thus the need for running all the filters in parallel becomes evident. Now, filter j is defined as the active filter if

$$e_j(k) < e_i(k). \tag{7}$$

The equation (7) provides a single-step decision logic for filter switch-over; consequently there is no delay and our algorithm is amenable to real time tracking. Using our proposed filter bank approach and single step decision logic to switch over among the filters, it is possible to track both maneuvering and non-maneuvering targets simultaneously.

4. SIMULATION RESULTS

In our simulations, we have used two filters: constant acceleration (CA) and Singers' maneuver model (SMM) for the multiple filter bank. For the simulations, the tracker is setup after object continuity is found in three consecutive frames in the sequence. The state parameters for each model used in tracking are initialized using the target position found in the first two frames of the sequence. We have evaluated the performance of the proposed algorithm using a number of sequences. Due to space limitation simulation results are presented only for three types of sequences: (a) clip 1 with one non-maneuvering approaching target, (b) clip 2 with three targets (one maneuvering and two non-maneuvering) and (c) clip 3 with four targets (two maneuvering and two non-maneuvering). For clip 2 and clip3, we have used synthetically generated IR target signature. For all the clips the trajectories are generated in the trajectory

Table 1. Performance Analysis of Track While Scan.

			Detection		Tracking	
Clip	A	B	C	D	E	F
1	1	40	0.97	0.17	1	0
2	3	70	0.99	0.18	3	0
3	4	70	0.93	0.07	4	0

A Number of targets
B Number of frames
C Detection rate = $\frac{\text{NC}}{\text{NC + NM}}$
D False alarm rate = $\frac{\text{NF}}{\text{NC + NF}}$
NC Number of correct detection
NM Number of mis detection
NF Number of false detection
E Number of successful tracks
F Number of false track (due to clutter)

plane placed at some depth in Z-direction. The trajectory plane is rotated about any of (X,Y,Z) axis with respect to the origin and then it is projected on to the image plane. Table 1 depicts the performance evaluation for the proposed track while scan algorithm. It is evident from the table that there is no false track established due to clutter though there are few false detections.

Figures 1(a) and 1(b) represent the output of the detection algorithm at frame 8 and frame 32 for clip 1. In Figure 1(c), the tracked trajectory is shown for the same clip. In the clip, clutter (bird) appears with the target for a number of frames. Our proposed method is able to discriminate between the target and the clutter. Note, due to space limitations, we have depicted the details of trajectories in clip 2 only. For clip 2, for all trajectories, initial target signature is of 3×3, gradually increasing to a maximum size of 15×11. The enlarged view of the target signature is shown in Figure 2(c). First trajectory is generated using constant velocity, coordinated turn and constant acceleration model. The initial position and initial velocity are set to $(9km, 7km)$ and $(280.0m/s, 280.0m/s)$. Trajectory takes total of four turns: (a) from frame 15 to 22 with turn angle 15°, it results in $\approx 10.5g$ acceleration, (b) from frame 30 to 36 with turn angle -15° ($\approx 10.27g$ acceleration), (c) from frame 43 to 52 with 12° ($\approx 8.8g$ acceleration) and (d) from frame 58 to 67 with -12° ($\approx 6.1g$ acceleration). For projecting the trajectory on to the image plane, the trajectory plane is assumed at 8km depth and rotated about (X,Y,Z) axis by $(15^\circ, 25^\circ, 12^\circ)$ respectively. Second trajectory is generated with initial position $(7km, 7km)$, initial velocity $(300m/s, 270m/s)$ and initial acceleration $(100.0m/s^2, 80.0m/s^2)$. The trajectory plane is assumed to be at 7km and it is rotated by 10° about X-axis and by 40° about Y-axis for projection on to the image plane. Finally, the third trajectory is generated with initial position $(10km, 7km)$ and initial velocity $(320m/s, -260m/s)$. For projecting the trajectory on to the image plane, the trajectory plane is rotated by 20° about X-axis and by 30° about Y-axis. Figures 2(a) and 2(b) represent the detection output at frame 20 and the tracked trajectory at frame 69. The condensation trail formed at the wing tips appears as clutter in the sequence.

Clip 3 is similar to clip 2 except for: (a) it has 4 targets, (b) maximum target size is 16×7. Figures 3(a) and 3(b) depict the detected target at frame 4, and the tracked trajectories. The enlarged view of the target signature is shown in Figure 3(c).

5. CONCLUSION

In the absence of any feature information other than motion, our wavelet based target detection algorithm is able to detect approaching targets in IR clips. Our multiple filter bank based tracking algorithm uses only two filters, namely, CA and SMM. It successfully tracks maneuvering and non-maneuvering targets under low contrast, clutter and continuously varying background. It does not use any apriori information about the target dynamics.

Acknowledgments

This research project was supported by Department of Extramural Research and Intellectual Property Rights, New Delhi, India.

6. REFERENCES

[1] G. Hewer et al., "Detection of small objects using adaptive wavelet-based template matching," in *Proc. SPIE Conference on Signal and Data Processing of Small Targets*, Oct. 1999, vol. 3809, pp. 95–106.

[2] Courtney I. Hilliard, "Selection of a clutter rejection algorithm for real-time target detection from an airborne platform," in *Proc. SPIE Conference on Signal and Data Processing of Small Targets*, July 2000, vol. 4048, pp. 74–84.

[3] Ulisses Braga-Neto, Manish Choudhary, and John Goutsias, "Automatic target detection and tracking in forward-looking infrared sequences using morphological connected operators," *Journal of Electronic Imaging*, (In press) 2004.

[4] Namrata Vaswani, Amit K Agrawal, Qinfen Zheng, and Rama Chellappa, "Moving object detection and compression in IR sequences," in *Computer Vision beyond the Visible Spectrum*, B. Bhanu and I. Pavlidis, Eds., chapter 5, pp. 153–177. Springer, 2003.

[5] Shao-Yi Chien et al., "Efficient moving object segmentation algorithm using background registration technique," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 12, no. 7, pp. 577–586, July 2002.

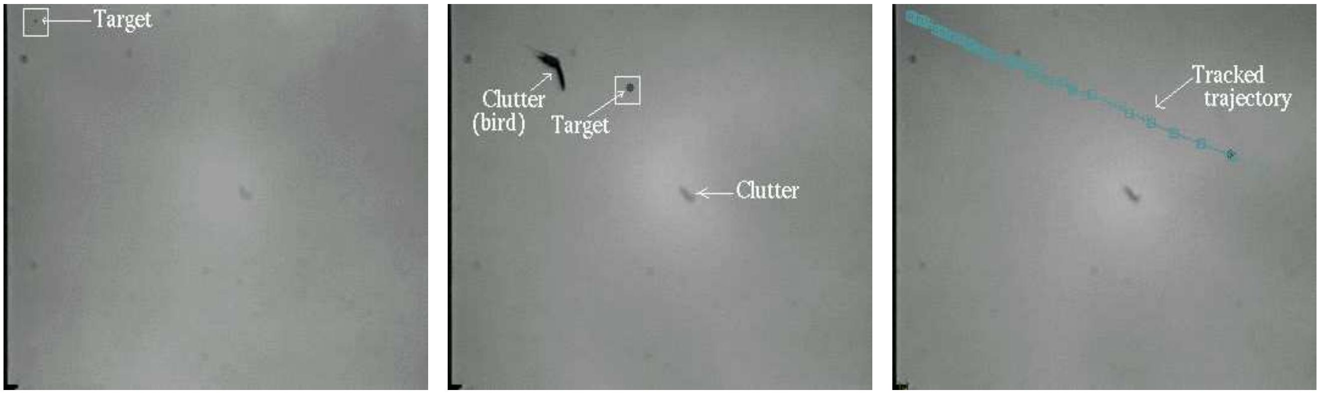
[6] Emrullah Durucan and Touradj Ebrahimi, "Change detection and background extraction by linear algebra," *Proceedings of IEEE*, vol. 89, no. 10, pp. 1368–1381, October 2001.

[7] Yaakov Tsaig and Amir Averbuch, "Automatic segmentation of moving objects in video sequences: A region labeling approach," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 12, no. 7, pp. 597–612, July 2002.

[8] Alper Yilmaz, Khurram Shafique, and Mubarak Shah, "Target tracking in airborne forward looking infrared imagery," *Image and Vision Computing, Elsevier*, vol. 21, no. 7, pp. 623–635, July 2003.

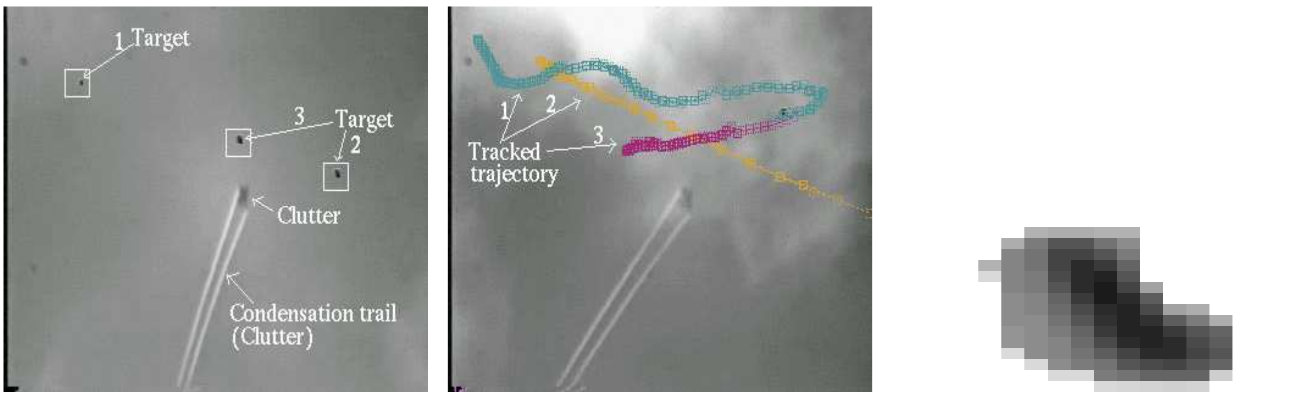
[9] Mukesh A. Zaveri, S.N. Merchant, and Uday B. Desai, "Multiple single pixel dim target detection in infrared image sequence," in *Proc. IEEE International Symposium on Circuits and Systems*, Bangkok, May 2003, pp. 380–383.

[10] Mukesh A. Zaveri, Uday B. Desai, and S.N. Merchant, "Tracking multiple maneuvering point targets using multiple filter bank in infrared image sequence," in *Proc. IEEE International Conference on Acoustics, Speech, & Signal Processing*, Hongkong, April 2003, pp. 409–412.



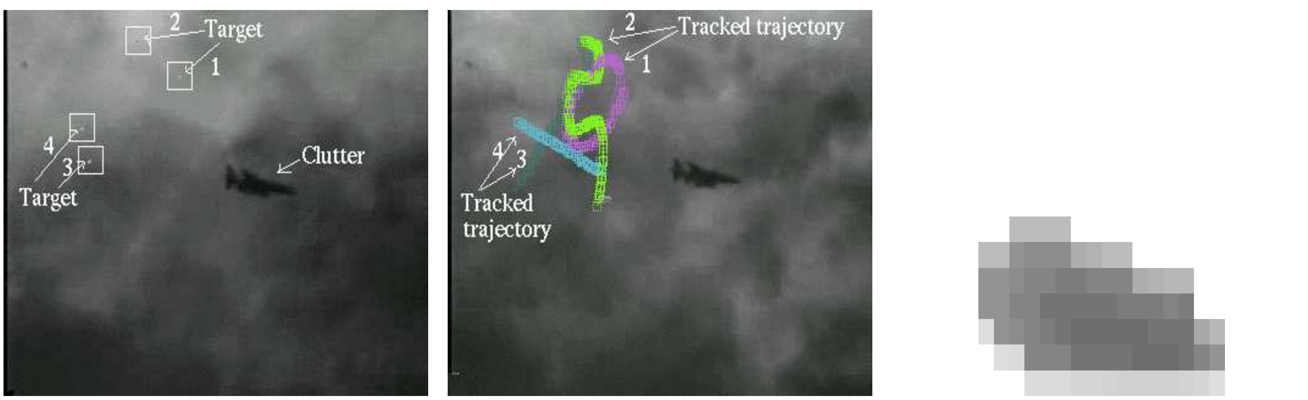
(a) at frame 8 (b) at frame 32 (c) trajectory plot at frame 39

Fig. 1. Approaching Target Detection and Tracking (clip 1)



(a) Detection at frame 20 (b) Trajectory Plot at frame 69 (c) IR target signature (Enlarged view)

Fig. 2. Approaching Target Detection and Tracking (clip 2)



(a) Detection at frame 4 (b) Trajectory Plot at frame 69 (c) IR target signature (Enlarged view)

Fig. 3. Approaching Target Detection and Tracking (clip 3)

[11] Suyog D. Deshpande, M.H.Er, V. Ronda, Philip Chan, “Max-Mean and Max-Median filters for detection of small-targets,” in *Proc. of SPIE, Signal and Data Processing of Small Targets*, Denver, Colorado, July 1999, vol. 3809, pp. 74–83.

[12] Jean-Michel LÉTANG, Patrick BOUTHEMY & Véronique REBUFFEL, “Robust motion detection with temporal decomposition and statistical regularization,” Tech. Rep., INRIA, Oct. 1995.

[13] Samuel S. Blackman, *Multiple-Target Tracking with Radar Applications*, Artech House, Inc., Boston, 1986.