Fast Detection of Multiple Point Targets in IR Sequence

Ankit Kulin, Sushil Kumar, U.B. Desai, S.N. Merchant
SPANN Laboratory
Electrical Engineering Department
Indian Institute of Technology, Bombay
Email: [akulin, sushilk, ubdesai, merchant]@ee.iitb.ac.in

Abstract—Detection of point targets in Infra-red (IR) video sequence based on selective median filtering is described here. The proposed algorithm is based on extraction of target information from the background using spatial filtering subsequently using a temporal filter for removing binary noise, cloud and clutter. The method utilizes the median filter to differentiate the potential targets in any region in the image and exclude the background. The results of simulation on real IR images indicate that the method can suppress strong cloud clutter as well as binary noise and efficiently detect small targets.

I. INTRODUCTION

Target detection and tracking play a significant role in surveillance applications. The primary purpose of IR sensor is to collect heat radiation from the sources in view and map them into intensity images. The images obtained thus are further processed to detect and track targets in them. The objective of search and track system is to detect targets and track them in real time.

The focus of the paper is to detect airborne targets which appear as points on the captured image. The typical features of the algorithm should include real time response, low false alarm rates and lower computational overhead.

Over the past few decades many algorithms have been proposed for the detection of small targets. The most commonly known detectors for this task use wavelet filters, image segmentation in watershed algorithm and spatial gradient approach.

The wavelet detection algorithm [1] uses temporal multi-scale decomposition for detecting point as well as small targets. However, the algorithm performs poorly when the targets are of low contrast. The algorithm is also computationally intensive to be implemented in real-time applications.

The watershed algorithm [2] is based on segmentation of objects in a scene, it doesn’t detect point targets.

The spatial gradient algorithm [3] performs favorably for point targets even with low contrast. However, its performance deteriorates in the case of blob targets with smoother edge gradation (intensity slowly merging) to the underlying cloud clutter.

This paper presents a new algorithm to suppress the strong clutter. The proposed spatial filter successfully enhances the dim targets (grey level difference 15) in the IR image and suppresses the surrounding clutter. It can be used for small/pixel sized target as well as large/blob targets (~30 pixels). The algorithm detects target with movement of 20-22 pixels between successive frames. The proposed algorithm has a very low computational overhead and is suitable for real time implementation. Temporal filtering using trajectory and intensity continuity in consecutive frames is used to discard the false targets and remove binary noise. Binary noise here refers to the salt and pepper noise of varying intensities present in the image due to errors introduced by sensor.

The algorithm is described in detail in Section II. Simulation results are reported in Section III followed by conclusions in Section IV.

II. DETECTION OF POINT/SMALL TARGETS

Point targets refer to the type of targets whose size ranges from a few pixels to one pixel, or even sub-pixel size in the image plane of a sensor. For single-pixel sized targets, conventional pattern recognition methods fail owing to lack of geometry and texture information. A noisy environment consisting of background clutter and possibly sensor noise presents the biggest obstacle of all. Randomly distributed high-intensity noise pixels would have the same appearance as the targets in a frame. In case of maneuvering targets, assumption of straight-line trajectory is in-valid. The case of multiple targets also poses a challenge.

A. The image model

The IR image sequence with small target embedded in the cloud clutter can be modeled as

\[ f(x, y, k) = S(x, y, k) + N(x, y, k) + C(x, y, k) \]  (1)

where, \( x \) and \( y \) represent the spatial coordinates of the acquired image with size \( N_x \times N_y \), \( k \) is the frame number, \( S(x, y, k) \), \( N(x, y, k) \), \( C(x, y, k) \) denote the target signal, the zero mean measurement noise, and clutter background respectively. The small target occupying only a few pixels is modeled as

\[ S(x, y, k) = \sum_{i=-k_m}^{k_m} \sum_{j=-k_m}^{k_m} A_{i,j,k} \delta(x-i, y-j, t-k) \]  (2)

where, \((2k_m + 1)x(2k_m + 1)\) is the size of target at time \( k \), \( \delta \) is two dimensional kronecker delta and \( A_{i,j,k} \) is an unknown signal intensity.
B. Selective median filter approach

In the proposed algorithm we construct an image \( I(x, y, k) \), of the background by removing the targets using a spatial filter. The spatial filter replaces the central pixel with a local median of intensities inside the window.

The background reconstruction is performed using a median filter of window size \( F \times F \) given by the equation:

\[
E = \text{median}(f(x + 2i, y + 2j, t - k))
\]

where \( i, j \in [-\text{floor}\left(\frac{F-1}{2}\right), \text{floor}\left(\frac{F-1}{2}\right)] \)

By considering alternate pixel position for the median filter we are able to capture even dim targets with gradually decreasing intensity. This outperforms the adjacent pixel filter in removing the clutter and it reduces the number of operations by half.

The cloud clutter removal procedure can be formulated as:

\[
\hat{f}(x, y, k) = |f(x, y, k) - I(x, y, k)|
\]

where \( f(x, y, k) \) is the original image, \( I(x, y, k) \) is the background image (calculated using the selective median filter), \( \hat{f}(x, y, k) \) is the residual image.

Now, we can construct the potential target map by using a threshold to enhance the targets. The threshold can be decided based on the minimum grey level difference the user needs to detect.

\[
D(x, y, k) = \begin{cases} f(x, y, k), & \text{if } \hat{f}(x, y, k) > \text{th} \\ 0, & \text{others} \end{cases}
\]

\( D(x, y, k) \) thus obtained is an intensity map of potential targets detected by the algorithm.

C. Elimination of false targets based on trajectory continuity

The residual image includes not only real targets but also binary noise. Noise is random and uncorrelated between consecutive frames, the target movement on the other hand is continuous and correlated. Therefore the small target’s trajectory is continuous. These characteristics, ensure that the target will appear in a small neighborhood in the following frame with high probability. The illumination of target in consecutive frames also remains constant. Thus we can define the conjunction function as follows.

\[
T(x, y, i, j) = |D(x + i, y + j, k + 1) - D(x, y, k)|
\]

\( i, j \in (-s, s) \)

where, \( s \) is a positive constant which denotes the search domain in successive frames. The decision for each pixel is obtained by thresholding the absolute difference of the intensity map \( T(x, y, i, j) \), using the following conjunction function.

\[
f(x, y, k) \in \text{target} \quad \text{if} \quad T(x, y, i, j) > \eta
\]

where, \( \eta \) is a positive constant.

III. Simulation results

For simulation, we use real images captured using IR sensor and synthetic images generated by IR scene simulator [7]. The clutter is mainly the strong undulant and bright cloud. Because the target is far away from the sensor, it appears as a tiny spot.

III. Simulation results

For simulation, we use real images captured using IR sensor and synthetic images generated by IR scene simulator [7]. The clutter is mainly the strong undulant and bright cloud. Because the target is far away from the sensor, it appears as a tiny spot.
background is 25-30 grey-levels for 8-bit image resolution with low signal to noise or signal to clutter ratio (SNR or SCR). For the simulation we choose window of size 9x9 for background reconstruction.

For the simulation we choose window of size 9x9 for background reconstruction.

The sequences have significant amount of binary noise. The results obtained after processing these IR images using proposed algorithm and applying temporal filtering on two frames are shown in Figure 2. The residual image has the potential targets only, almost all of the clutter and background has been removed in the process. For testing the algorithm on point targets with a grey-level difference of 15-20 from the background, we use an image sequence artificially generated by the IR scene simulator. A frame from this sequence is shown in Figure 3.

The image is then embedded with 0.5% binary noise with varying intensity and pixel position to simulate sensor noise as shown in Figure 4(a). On testing the proposed algorithm, we first obtain cloud clutter suppressed image Figure 4(b). On using temporal filtering using two successive frames we observed that all of the binary noise gets removed and we are left with the two targets as depicted in Figure 4(c).

In our simulation, 4 image sequences are used for testing the proposed algorithm. Table I tabulates the detection accuracy of the proposed algorithm. The results indicate a higher accuracy of the proposed algorithm compared to that of other algorithms.

The comparison of the simulation results of the proposed algorithm with three other algorithms [1,2,3] on the four sequences described earlier are tabulated in Table II. The wavelet and watershed based detection algorithms fail to detect point targets in case of synthetic and embedded image sequences.

### TABLE I

<table>
<thead>
<tr>
<th>Image sequence</th>
<th>No. of targets</th>
<th>Correct detection</th>
<th>Incorrect detection</th>
<th>Detection Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>synthetic</td>
<td>90</td>
<td>90</td>
<td>0</td>
<td>100%</td>
</tr>
<tr>
<td>sequence1</td>
<td>400</td>
<td>400</td>
<td>0</td>
<td>100%</td>
</tr>
<tr>
<td>sequence2</td>
<td>600</td>
<td>580</td>
<td>20</td>
<td>96.66%</td>
</tr>
<tr>
<td>sequence3</td>
<td>60</td>
<td>60</td>
<td>0</td>
<td>100%</td>
</tr>
</tbody>
</table>

### TABLE II

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Synthetic (90)*</th>
<th>Sequence1 (400)*</th>
<th>Sequence2 (600)*</th>
<th>Sequence3 (60)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed</td>
<td>90</td>
<td>400</td>
<td>580</td>
<td>60</td>
</tr>
<tr>
<td>Wavelet filter[1]</td>
<td>-</td>
<td>340</td>
<td>552</td>
<td>-</td>
</tr>
<tr>
<td>Watershed[2]</td>
<td>-</td>
<td>260</td>
<td>408</td>
<td>-</td>
</tr>
<tr>
<td>Spatial gradient [3]</td>
<td>90</td>
<td>51</td>
<td>332</td>
<td>126</td>
</tr>
</tbody>
</table>

* denotes the number of targets in sequence
- denotes no detection

### IV. CONCLUSION

Detection and tracking of point targets present several unique challenges. Features like clutter, multiple targets, maneuvering targets, absence of unique signature and false alarm
The proposed algorithm uses selective median filter to reconstruct the cloud clutter and subtract it from the initial image. The median filter performance is evident from the residual image, which almost has zero background or clutter. It is observed that the proposed algorithm performs efficiently with the point as well as blob targets, even when the background is highly evolving and the targets are of low contrast. The algorithm can be implemented real time.

References


