MULTIPLE SINGLE PIXEL DIM TARGET DETECTION 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

It is difficult to detect and classify multiple point targets in the presence of clutter and noise, in air borne IR image sequence. Most existing approaches assume targets of several pixels size or of Gaussian shape with variance of 1.5 pixel. We propose a wavelet based detection technique for multiple single pixel dim targets in IR image sequence in the presence of clutter and noise. Temporal multiscale decomposition is used to detect temporal change. This temporal change map is postprocessed to reduce probability of false alarm and misdetection.

1. INTRODUCTION

Most of IRST (Infrared Search and Track) systems assume Gaussian distributed size or blob size target [1]. In such cases, matched filtering [12], [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 [14], [5]) and exploitation of motion.

Adaptive algorithm based on TDMS has been proposed for target detection in [6]. In our simulations, we found that TDMS algorithm is very slow in case of large frame size. Recently, nonlinear filters have been widely used for clutter suppression [7], [8]. Sequential probability ratio test [9] and dynamic programming [10] based approaches proposed by Blostein and Barniv are computationally expensive. 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 [11]. For naval surveillance, a method is proposed in [12] 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 [13].

In our proposed method we apply temporal decomposition to detect multiple targets in the presence of clutter and noise, using change detection map. Change detection map is postprocessed to reduce false alarm and misdetection. The paper is organized as follows: temporal multiscale decomposition is presented in Section 2.1 which builds change detection map. It is postprocessed to identify the targets, and this is described in Section 2.2. Simulation results are presented in Section 3. It shows that our proposed method is able to detect dim targets in a real as well as synthetic images with signal to clutter plus noise ratio > 2 dB.

2. WAVELET BASED TARGET 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. 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 gives 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 \( \psi_{mn}(t) \) can be written as follows:

\[
\psi_{mn}(t) = a^{-m/2} (a^{-m} t - nb)
\]

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 \( C^{k+1} \)

\[
C^{k+1}(i) = \sum_{n} H(n-2i)C^k(n)
\]

\((H \text{ is equivalent to a low pass filter})\)

(3)

\[
D^{k+1}(i) = \sum_{n} G(n-2i)C^k(n)
\]

\((G \text{ is equivalent to high pass filter})\)

\( D^1 \) characterizes high temporal frequencies components. Following levels \( D^2, D^3 \ldots \) correspond to lower frequency bands. The Hart basis is used in our experimentation. This is because, with an increase in the number of wavelet filter coefficients, larger number of image frames are required for temporal multiscale decomposition, this in turn introduces larger delay in making the decision.

The advantage of temporal multiscale decomposition is that no preprocessing technique, like spatial smoothing (low pass filter or median filter) is required. For single pixel targets, preprocessing can have a disastrous effect, because it may completely
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)

eliminate the target. 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. 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:

$$H_0: \quad \psi^k(p) \geq \lambda$$

$$H_1: \quad \psi^k(p) < \lambda$$

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

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

and follows a $\chi^2$ distribution with three degrees of freedom. $\lambda$ is a threshold which may be inferred from tables of statistical laws. $N$ 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 with respect to the centre of the window. $\sigma^2$ is the variance of the pixel intensity within the window. At least three levels of the wavelet transform are required to correctly discriminate the 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 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.
- If a pixel at time $t$ is detected as temporal change 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 $\delta_k$ 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 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

$$lc(x_n, y_n) = \left| I(x_n, y_n) - \frac{1}{s_1} \sum_{(x_m, y_m) \in A_f} I(x_m, y_m) \right|$$

where $I(x_n, y_n)$ is the gray level value of the pixel at $(x_n, y_n)$, $s_1$ is the size of neighborhood window in terms of pixels and $A_f$ is defined as

$$A_f = \{(x_j, y_j) | (x_j, y_j) \in N_i(x_n, y_n) \land (x_n, y_n) \neq (x_j, y_j)\}$$
where \( N_r \) is the neighborhood window defined by a circle of radius \( r \) centered at \((x_n, y_n)\). \( I_c(x_n, y_n) \) is compared with predefined threshold \( \rho \). If it crosses the threshold it may be a point target. This 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 the intensity of pixels within eight-connected region only. Consider a pixel at \((x_n, y_n)\), we accept it as a candidate target: if \(| I(x_n, y_n) - I(x_m, y_m) | < \epsilon \) 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.

3. SIMULATION AND RESULTS

Synthetic Infrared images are generated using real time temperature data [14]. 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 \(1024 \times 256\) and the single pixel target movement is \(\pm 20\) pixels per frame. In simulation, eight frames are used to detect the target. The values of various parameters chosen for the simulation are as follows: \(\lambda = 13.277\) (with degree of freedom = 3), \(\Delta\theta = 0, r = 4, \rho = 30\), \(\epsilon = 10\). 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. The change detection map for another IR video clip with two dim targets is shown in Figure 3 and 4 for frame number 26 and 38 respectively. The target identified after postprocessing can be seen in Figure 5 and 6 for frame number 26 and 38 respectively. In frame number 38, one of the targets has intensity value comparable with that of cloud and in such a case, the algorithm fails to detect temporal change at this position (misdetection marked with an ellipse). We have also tested the proposed target detection algorithm on real IR image sequence (Figure 7). For simulation, a point target is embedded in IR image frames. In a frame 29 of real IR image sequence, false alarm occurs due to flare (marked as FA). We define the signal to clutter + noise ratio (SCNR) as

\[
SCNR = 10 \log \left( \frac{S_t - m_0}{\sigma_0^2} \right)
\]  

\(S_t\) - measured signal intensity 
\(m_0\) - mean of the background noise 
\(\sigma_0^2\) - variance of the background noise
Figure 6: Target Identification after Postprocessing of Change Detection Map at frame number 38 (IR clips:2)

Figure 7: Frames for real IR image sequence

where $S_1$ is the minimum intensity value at a pixel in the presence of target, $m_3$ is the average value of the clutter plus noise and $\sigma_2^2$ is clutter plus noise power. With a proper choice of $\varepsilon$ and $\delta_{th}$ the proposed scheme works very well with SCNR $\geq$ 2 dB.

4. CONCLUSION

For large frame size and large target movement, wavelet based detection scheme performs extremely well in the presence of clutter and noise. Performance of wavelet based target detection scheme degrades for highly maneuvering target and for targets with more than 30 pixels difference between positions in two consecutive frames.

5. REFERENCES


