Point Target Detection using Wavelet in InfraRed image sequence

MUKESH A. ZAVERI  UDAY B. DESAI  S. N. MERCHANT
SPANN Lab, Electrical Engineering Dept., IIT Bombay - 400076. INDIA

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 algorithm for single pixel targets. The proposed scheme works very well with SCNR ≥ 2 dB.

Key-words:- Wavelet, Temporal Multiscale Decomposition, Signal to Clutter plus Noise Ratio (SCNR)

1 Introduction

Most of IRST (Infrared Search and Track) systems assume Gaussian distributed size or blob size target [1]. In such cases, matched filtering ([2], [3]) 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 ([4], [5]) and exploitation of motion.

Adaptive algorithm based on TDLMS (Two Dimensional Least Mean Square) is proposed for target detection in [6]. 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 ([7], [8]). Sequential probability ratio test [9] and dynamic programming (DPA) [10] based approaches proposed by Blostein and Barniv are computationally expensive. In [11] 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 [12] 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 [13]. For naval surveillance, a method is proposed in [14] 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 [15].

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 may be given to the tracking phase as candidate target list.

Our algorithm for point target detection is as follows: the first phase gives binary map i.e. either temporal change detected or not and second phase performs postprocessing to reduce false alarm or misdetection. The paper is organized as follows. In Section 2.1, temporal decomposition is introduced to obtain change detection map. These maps are postprocessed by the method described in section 2.2. Simulation results are presented in Section 3.

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 multi-
 scale decomposition.

2.1 Temporal 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 tem-
poral integration is required in absence of any texture
information. All these temporal signals can be con-
sidered 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 ba-
sis is composed of a family of functions adjusted by
two parameters: one for the position (in time) and the
other for the scale, . 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 \( C^{k+1} \):

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

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

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

(\( G \) is equivalent to high pass filter). \( D^1 \) charac-
terizes 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 tem-
poral multiscale decomposition, which introduces fur-
ther delay in making decision. The advantage of tem-
poral multiscale decomposition is that no preprocessing
 technique, like spatial smoothing (low pass filter or
median filter) is required. Preprocessing has the dis-
advantage 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 tempo-
ral change or not. These binary maps are interme-
diate 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 hy-
potheses are compared: hypothesis \( H_0 \) (no temporal
change at \( s \)) and hypothesis \( H_1 \) (temporal change at \( s \)).
The log-likelihood ratio corresponding to hypotheses
\( H_1 \) and \( H_0 \) is derived and decision step is formalized as:

\[
\begin{align*}
H_0 & \quad \psi^k(p) < \lambda \\
H_1 & \quad \psi^k(p) \geq \lambda
\end{align*}
\]

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

\[
\begin{align*}
\psi^k(p) &= \frac{1}{2\sigma^2_k}\left[ \frac{1}{N}\left( \sum_{i=1}^{N} X_i^2 \right) - \frac{1}{N}\left( \sum_{i=1}^{N} X_i \right)^2 \right] \\
& \quad + \frac{1}{2}\left( \sum_{i=1}^{N} \left( \frac{1}{N}\left( \sum_{i=1}^{N} Y_i \right)^2 \right) - \frac{1}{N}\left( \sum_{i=1}^{N} Y_i \right)^2 \right)
\end{align*}
\]

\( \psi^k(s) \) follows \( \chi^2 \) 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 lo-
action of pixels w.r.t. the centre of window. \( \sigma_k^2 \) is vari-
ance of the pixel intensity within the window. At least
three levels of the Wavelet transform is required to al-
low us to correctly discriminate the various dynamical
behavior present in the scene. The following heuris-
tics 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.
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 Change Detection Map

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 \( \delta_{th} \) 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 \( B_t \) of the segmented image. \( lc(x_n, y_n) \) is defined as

   \[
   lc(x_n, y_n) = |I(x_n, y_n) - \frac{1}{s_t} \sum_{(x_m, y_m) \in A_f} I(x_m, y_m)|
   \]  

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

   \[
   A_f = \{(x_j, y_j) | (x_j, y_j) \in N_t(x_n, y_n) | (x_n, y_n) \neq (x_j, y_j)\}
   \]

   with \( N_t(x_n, y_n) \) the neighborhood of size \( t \).
where $N_r$ is the neighborhood window defined by a circle of radius $r$ centered at $(x_n, y_n)$. $l(x, y)$ is compared with predefined threshold $\rho$. 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 $\forall (x_m, y_m) \in N_8(x_n, y_n), m \neq n$, where $N_8(x_n, y_n)$ 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 [16]. 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 target movement is $\pm 20$ pixels per frame. 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 define the signal to clutter + noise ratio (SCNR) as

$$SCNR = 10 \log \left( \frac{(S_t - m_0)^2}{\sigma_0^2} \right)$$  \(7\)

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 $\sigma_0^2$ is clutter plus noise power. With a proper choice of $\varepsilon$ and $\delta_h$ the proposed scheme works very well with $SCNR \geq 2$ dB.

4 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 technique is satisfactory even in the presence of multiple targets.

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


