Track before detect for point targets with particle filter in infrared image sequences

Hongtao Hu (胡洪涛), Zhongliang Jing (敬忠良), and Shiqiang Hu (胡士强)

Institute of Aerospace Information and Control, School of Electric, Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200030

Received November 2, 2004

The problem of detecting and tracking point targets in a sequence of infrared images with very low signal-to-noise ratio (SNR) is investigated in this paper. A track before detect algorithm for infrared (IR) point target is developed based on particle filter. The particle filter is used to estimate the state of the target in track stage. The unnormalized weights of the output of the filter are used to approximately construct the likelihood ratio for hypothesis test in detection stage. Experiment results with the real image sequences that SNR is about 2.0 show that the proposed algorithm can successfully detect and track point target.

OCIS codes: 110.2970, 110.3080, 100.0100.

Detection and tracking of dim moving small targets in low signal-to-noise ratio (SNR) environments is a very important issue and difficult problem in infrared (IR) systems. Since image SNR is very low and target features like size, shape, texture are unavailable, small targets cannot be detected reliably using one image frame. Many researches have employed image sequences in the detection procedure to enhance the performance, which integrates the target energy over multiple frames. Among of them track-before-detect (TBD) approach is a powerful technique for point targets detection in low SNR image sequences. TBD does not perform target detection using single threshold at each measurement frame, but process data over a number of frames before decisions on target existence are made.

Over the last twenty years, many examples of TBD are available in the literature. In Ref. [1] three-dimensional (3D) matched filtering is applied to moving target detection using optical images. This method assumes that the target velocity is known. In Ref. [2] sequential probability ratio test (SPRT) is used to perform constant velocity target detection in image sequences. The target amplitude has been assumed constant, and the sensor point spread function has not been considered. In Ref. [3] dynamic programming (DP)-based TBD methods have been proposed. These methods avoid the problem of velocity mismatch, and can handle targets with slow maneuvers. However if the SNR is sufficiently low, these methods do not improve detection performance no matter how many frames of data are processed. Although the above-mentioned algorithms have good detection performance, their tracking performance is bad. If we want to get better tracking performance, some filter algorithms such as Kalman filter are needed. In Ref. [4] the TBD problem is approached via a Bayesian particle filter and performed on radar signal processing for tracking weak targets. But the detection performance is not considered.

In this paper, a TBD algorithm based on particle filter is performed on IR image sequences. This approach uses particle filter to estimate the posteriori probability density of a possible target in IR image. The output of the running particle filter, the unnormalized weight, is used to approximately construct the likelihood ratio for hypothesis test in detection stage. Compared with many previous methods this approach has a number of potential advantages: i) it can handle non-linear and/or non-Gaussian problem efficiently so that we can get good tracking performance; ii) it can avoid the problem of velocity mismatch so that we can get good detection performance; iii) the effects of a point spread function can be easily accommodated.

Assume that in a short time period, the target dynamics can be approximated by a straight-line track with constant velocity. We model the target motion by the white-noise acceleration model

$$X(k+1) = FX(k) + Gw(k), \quad (1)$$

$$X(k) = \begin{bmatrix} x(k) \\ \dot{x}(k) \\ y(k) \\ \dot{y}(k) \end{bmatrix}, \quad F = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$G = \begin{bmatrix} T^2/2 & 0 \\ T & 0 \\ 0 & T^2/2 \\ 0 & T \end{bmatrix},$$

where $(x(k), y(k))$ and $(\dot{x}(k), \dot{y}(k))$ are the target pixel positions and velocities, respectively. T is the time between successive sensor readings and termed frames. $w(k)$ is zero-mean, mutually independent, white Gaussian noise sequences with covariance matrices $Q(k)$.

Consider a model of image ($M \times N$ pixels) data measurements with additive noise

$$z_k(i,j) = s_k(i,j) + b_k(i,j) + v_k(i,j),$$
$$1 \le i \le M, 1 \le j \le N, \quad (2)$$

where $z_k(i,j)$ is the intensity of pixel (i,j); $s_k(i,j)$ is intensity affected by the target in pixel (i,j); $b_k(i,j)$ is the intensity of a deterministic clutter background, and $v_k(i,j)$ is measurement noise.

Assume that there may be only one target in the scene,

$s_k(\cdot)$ is modeled as

$$s_k(i,j) = A(k)h(i,j,x(k),y(k)), \tag{3}$$

where $h(\cdot)$ is the sensor point spread function (PSF); $A(k)$ and $(x(k),y(k))$ are target peak intensity and pixel position at moment k, respectively. For long-range infrared search and track system (IRST) applications, the target can be modeled as a point source. A truncated 2D Gaussian density with circular symmetry is a common model for $h(i,j,x,y)$[5]. That is

$$h(i,j,x,y) = \begin{cases} \frac{1}{2\pi\sigma_x^2\sigma_y^2}\exp\left(-\frac{(i-x)^2}{2\sigma_x^2}-\frac{(j-y)^2}{2\sigma_y^2}\right) & \text{for } |i-x|<3 \text{ and } |j-y|<3 \\ 0 & \text{otherwise} \end{cases}, \tag{4}$$

where $\sigma_x^2 = \sigma_y^2 = 1$. $b_k(i,j)$ can be removed through some clutter removal algorithms. $v_k(i,j)$ is assumed to be zero mean Gaussian noise. The variance is estimated as

$$\sigma_v^2(i,j,k) = \frac{1}{L}\sum_{a=k-L+1}^{k}[z_k(i,j)-\hat{\mu}_k(i,j)]^2, \tag{5}$$

where $\hat{\mu}_k(i,j)$ is the intensity of estimated background, L is the number of frames processed. For reducing the effect of the target, a 2D median filter is used.

The dynamical system for the target is composed of Eqs. (1) and (4). Particle filter is a new filtering method based on Bayesian estimation and Monte Carlo method and can effectively cope with complicated nonlinear and/or non-Gaussian problems. For a basic particle filter, at each instant k, a new set of particles is drawn from the Markovian transition kernel $p(X(k)|X(k-1))$ according to Eq. (1), and associated importance weights are updated using the likelihood function $p(Z_k|X(k))$ according to Eq. (2). A selection step, consisting of resampling from the particle set with replacement according to the importance weights, is added to prevent the distributions increase. From the weighted particle set, we can then compute an estimate of the target state using, for example, a minimum mean-square error (MMSE) or a maximum a posteriori (MAP) criterion. A basic particle filter algorithm is presented as follows:

1) Set $k=0$ and generate N samples $\{X^{(i)}(0)\}_{i=1}^N$ from the initial distribution $p(X(0))$.

2) Compute the weights $w_k^i = w_{k-1}^i p(Z_k|X^i(k))$ and normalize, ie., $\tilde{w}_k^i = w_k^i / \sum_{j=1}^N w_k^j$, $i=1,2,\cdots,N$.

3) If resampling is applied, then generate a new set $\left\{\tilde{X}^i(k)\right\}_{i=1}^N$ by resampling with replacement N times from $\{X^i(k)\}_{i=1}^N$, with probability $\Pr\left\{\tilde{X}^i(k) = X^j(k)\right\} = \tilde{w}_k^j$ and $w_k^i = \tilde{w}_k^i = 1/N$; otherwise let $\left\{\tilde{X}^i(k)\right\} = \{X^i(k)\}$.

4) Output a set of particles $\left\{\left(\tilde{X}^i(k),\tilde{w}_k^i\right)\right\}_{i=1}^N$ approximating the posterior distribution. Expectation is $\hat{X}(k) = \sum_{i=1}^N \tilde{w}_k^i \tilde{X}^i(k)$; covariance is $P(k) = \sum_{i=1}^N \tilde{w}_k^i (\tilde{X}^i(k)-\hat{X}(k))(\tilde{X}^i(k)-\hat{X}(k))^T$.

5) Increase k and iterate to step 2).

As mentioned above, $p(X(k)|X(k-1))$ and $p(Z_k|X(k))$ are two important functions for particle filter. In the IR image sequences, according to Eq. (1) the Markovian transition kernel is given by

$$p(X(k)|X(k-1)) = p_w = N(FX(k-1), GQG^T). \tag{6}$$

The likelihood $p(Z_k|X(k))$ is given by[5]

$$p(Z_k|X(k)) = \prod_{i,j\in C(X(k))} p_{s+n}(z_k(i,j)|x,y) \prod_{i,j\notin C(X(k))} p_n(z_k(i,j)), \tag{7}$$

where $p_{s+n}(z_k(i,j)|x,y)$ is the probability density function (PDF) of the target signal + noise in pixel (i,j) given that the target is located at (x,y), $p_n(z_k(i,j))$ is the PDF of the sensor noise in pixel (i,j), $C(X(k))$ is the set of pixels affected by the target.

In particle filter, the weight for a particle is proportional to its likelihood. Since the weights are only required up to proportionality, we may divide by $\prod_{i,j} p_n(z_k(i,j))$ and get

$$w^i \propto \prod_{i,j\in C(X^i(k))} l(z_k(i,j)|x^i,y^i), \tag{8}$$

where $l(z(i,j)|x^i,y^i) = \frac{p_{s+n}(z_k(i,j)|x^i,y^i)}{p_n(z_k(i,j))}$.

For detection we use a likelihood ratio test (LRT) algorithm. The detection problem consists in deciding between the two hypotheses:

H_0: no target present

$$z_k(i,j) = v_k(i,j), \tag{9}$$

H_1: target present

$$z_k(i,j) = s_k(i,j) + v_k(i,j). \tag{10}$$

The likelihood ratio is defined by

$$L(Z_k,\cdots,Z_{k+l}) = \frac{p(Z_k,\cdots,Z_{k+l}|H_1)}{p(Z_k,\cdots,Z_{k+l}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda. \tag{11}$$

The choice of a good threshold λ is a kind of compromise between false alarms and the probability of detection. If we accept the hypothesis H_1, it means the target is detected and the state of target is the output of particle filter. Otherwise, we use the one step prediction of the filter as the state of target.

Combined with track stage introduced above, making reference to Ref. [4], we can compute likelihood ratio using

$$L(Z_k,\cdots,Z_{k+l}) = \frac{\prod_{j=0}^{l}\left(\sum_{i=1}^{N} w^i(j)\right)}{N^{l+1}}. \tag{12}$$

The brief proof of Eq. (12) is given below.

Proof: first observe that

$$p(Z_k,\cdots,Z_{k+l}|H_0)=\prod_{j=0}^{l}p_n(Z_{k+j}). \tag{13}$$

Under the hypothesis of a target being present, hypothesis H_1, we can write [dropping the notation $(\cdot|H_1)$]

$$p(Z_k,\cdots,Z_{k+l})=\prod_{j=0}^{l}p(Z_{k+j}|Z_{k+j-1}), \tag{14}$$

where

$$\begin{aligned}&p(Z_{k+j}|Z_{k+j-1})=\\&\int_{X(k+j)}p(Z_{k+j}|X(k+j),Z_{k+j-1})p(X(k+j)|Z_{k+j-1})\\&=E_{p(X(k+j)|Z_{k+j-1})}p(Z_{k+j}|X(k+j),Z_{k+j-1})\\&\approx\frac{1}{N}\sum_{i=1}^{N}p_{s+n}(Z_{k+j}|X^i(k+j)).\end{aligned}$$

Then the likelihood ratio can be rewritten as

$$\begin{aligned}L(Z_k,\cdots,Z_{k+l})&=\frac{\prod_{j=0}^{l}\left(\frac{1}{N}\sum_{i=1}^{N}p(Z_{k+j}|X^i(k+j))\right)}{\prod_{j=0}^{l}p_v(Z_{k+j})}\\&=\frac{1}{N^{l+1}}\prod_{j=0}^{l}\sum_{i=1}^{N}\left(\frac{p(Z_{k+j}|X^i(k+j))}{p_v(Z_{k+j})}\right)\\&=\frac{\prod_{j=0}^{l}\left(\sum_{i=1}^{N}w^i(j)\right)}{N^{l+1}},\end{aligned} \tag{15}$$

end.

Next we compare the tracking and detection performances of our proposed algorithm (PF-TBD) and the classic algorithm mentioned in Ref. [1] (MF) using a real IR image sequence that contains a certain type of aircraft at a distance of about 40 kilometers. The data consist of 50 frames of IR data (256×256 pixels in each frame). Since the frame rate is high, it is reasonable to consider that the intensity of target is constant in short time.

First, the method mentioned in Refs. [6,7] is used to suppress clutter. After image preprocessing a Monte-Carlo experiment has been performed with 100 replications.

In every experiment the initial prior PDF of its position is assumed to be uniform over the interesting region that is got through manual initialization or automatic initialization using a detection algorithm depending on the prior knowledge. The initial prior PDF of its velocity is assumed to be uniform over $[-0.3, 0.3]$ in the x and y directions. We use $N=3000$ particle numbers in particle filter. To alleviate the problem of sample degeneracy, we set $Q=\begin{bmatrix}0.05^2 & 0\\ 0 & 0.05^2\end{bmatrix}$. In the detection stage, we set slide-windows $L=4$. An easily calculated threshold λ as a function of the desired probability of false alarm is not possible. For simplicity, we let $\lambda=0.003$ which means that a prior probability ratio of the target presenting to the target not presenting is 1000:3.

Figure 1(a) shows the typical IR scenario investigated in this work. Most of the scene is occupied by the cloud so that the SNR is very low (about 2.0). Figure 1(b) shows the 40th frame detection result of once experiment. The target is marked by "+" in figure. Figure 2 shows the RMSE (relative mean square error) in pixels of the target position estimates, respectively in the vertical and horizontal directions. From Fig. 2 we can see that PF-TBD method quickly acquired the target after an initial error and tracked it with a low RMSE. While the MF method has a large error, which occurs because the mismatch between the actual motion model and the approximate discrete-state model. The PF-TBD method can estimate the target motion in real time so that it can reduce the mismatch and get a good tracking performance. The PF-TBD method can effectively calculate the likelihood ratio in detection stage and provide good tracking performance, which leads to a good detection performance. The Table 1 reflects the detection performance. From Fig. 2 and Table 1 we can see our proposed algorithm has a better performance both tracking and detection.

In this paper, a particle filter for TBD in IR image sequences is investigated. In track stage, a particle filter has been used to estimate the target state. In the detection stage, likelihood ratio test has been used and the likelihood ratio can be expressed in terms of the output of the particle filter, i.e. the unnormalized particle weights. From experiment results the proposed algorithm can

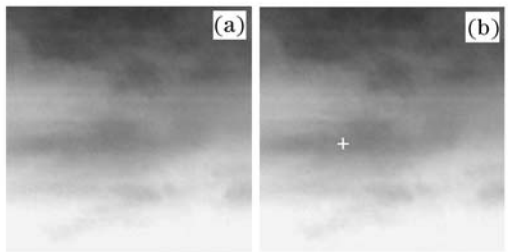

Fig. 1. Origin image and detection result: (a) origin image; (b) detection result.

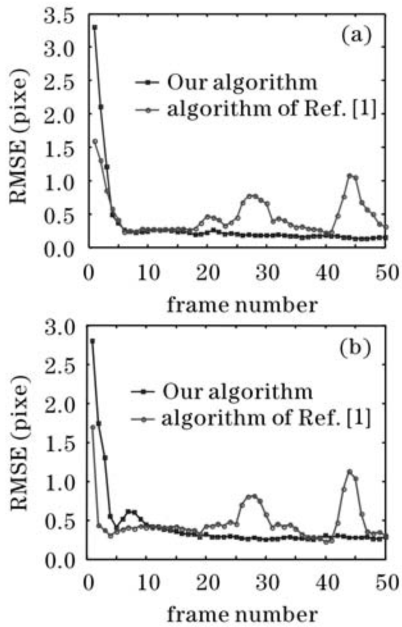

Fig. 2. RMSE of target pixel position: (a) horizontal direction; (b) vertical direction.

Table 1. Detection Performance

	Detection Probability	Missing Probability
Algorithm of Ref. [1]	80.3%	19.7%
Our Algorithm	96.7%	3.3%

effectively track and detect a target in low SNR IR image sequences.

This work was jointly supported by National Natural Science Foundation of China (No. 60375008), China PH.D Discipline Special Foundation (No. 20020248029), China Aviation Science Foundation (No. 02D57003), Aerospace Supporting Technology Foundation (No. 2003-1.3 02), EXPO Technologies Special Project of National Key Technologies R&D Programme (No. 2004BA908B07), Shanghai Key Technologies Preresearch Project (No. 035115009). H. Hu's e-mail address is hht@sjtu.edu.cn.

References

1. I. S. Reed, R. M. Gagliardi, and L. B. Stotts, IEEE Trans. AES **26,** 434 (1990).
2. S. D. Blostein and H. S. Richardson, IEEE Trans. AES **30,** 197 (1994).
3. L. A. Johnston and V. Krishnamurthy, IEEE Trans. AES **38,** 228 (2002).
4. Y. Boers and H. Driessen, IEEE SP Lett. **10,** 300 (2003).
5. D. J. Salmond and H. Birch, in *Proceedings of American Control Conference* **3755** (2001).
6. S. Leonov, IEEE Trans. AES **37,** 832 (2001).
7. L.H. Jiang, D. M. Ren, and X. H. Lu, Chin. Opt. Lett. **1,** 689 (2003).