Signature-Driven Multi-Target Tracking

Jian Kang Wu, Shuyan Sun, Sheng Jiang, Xiaoyi Ren, Zhipei Huang
School of Information Science and Engineering
Graduate University of Chinese Academy of Sciences
jkwu@gucas.ac.cn, sunshuy09b@mails.gucas.ac.cn

***Abstract** - Tracking multiple maneuvering targets remains a challenge because of clutter and spurious targets. We propose a Signature-Driven multiple target Tracking (SDT) method which fuses target data in spectral, spatial and temporary spaces to form signatures of targets. Markov properties of target features and dynamics are well defined in the signature of targets so that the data association process in SDT is efficient and effective. The experimental results have shown its superior performance.*

Keywords: Tracking, data association, Kalman filtering, sensor data fusion.

1 Introduction

Data association for multiple target tracking has been a challenging research topic especially for large amount of maneuvering target in a clutter situation, where possible data association hypotheses are huge. So far the joint probabilistic data association proposed by Bar-shalom [1] in 1974 and the Multiple Hypothesis Tracking (MHT) proposed by Reid [2] are two representative methods. After a lot of improvement [3, 4], MHT has been regarded as the best with respect to performance. The fundamental idea of MHT is to make use of the temporal Markov property of target data and makes decision after N-scan. However, it is based on hypothesis-test principle. The performance becomes poor when the number of targets and amount of clutter increase.

Different from the traditional target tracking research, in the area of visual tracking of multiple targets research has been focused on visual features of targets. Data association can be modeled as networks. By using visual features and spatial distances, cost can be assigned to each link between consecutive data points, and therefore, a linear programming or Viterbi search can be used to find optimal track [5, 6].

Traditional multiple target tracking mainly work on radar data on which targets appear as random points, while in visual tracking data of targets are within image frames. As advances of data acquisition techniques, in many applications data are acquired using multiple apparatus of different types and at various locations. We are provided with multiple spectral data, and data fusion in multiple targets tracking is invertible. We present here a Signature- Driven multiple target Tracking (SDT) method which fuses the spectral, spatial and temporary features of targets to form signature of targets based on the models and knowledge of target properties and dynamics. The overall belief of a target signature is evaluated based on the Markov property of target dynamics, consistency of consecutive spatial features, and the detection probability. This overall belief of the signature is then used to reduce the number of hypotheses at the earliest time before the filtering process, and therefore to increase the accuracy and reduce the computation complexity of the tracking algorithm.

The novelties of SDT are, first, instead of track confirmation and termination using track score in MHT, we introduced a concept of target signature, which is the model-based fusion of spectral, spatial and temporal data. It provides the most comprehensive information for the target detection and tracking. Second, in the target signature, we derive Markov properties for both target dynamics, and target spectral and spatial features. As a matter of fact, it is combination of advantages of tracking methods in both traditional target tracking and visual tracking. Thirdly, as far as process concerned, instead of filtering-and-track confirmation in MHT, SDT performs signature formation and evaluation in N-scan, and then the most likely signature is retained and sent for filtering. This will reduce the chance of clutters going to the filter so that target tracks after filtering have high accuracy.

In what follows, section 2 provides the background and notations for multiple target tracking, SDT method is described in Section 3. The experimental results are presented in section 4. Finally, section V is the conclusion.

2 Background and Notations

The multiple target tracking is based on single target tracking together with data association. In single target tracking, the state of a target here is assumed to be a first order Markov process on the state space $\chi \in R^{nx}$ with transition density $p(x_t|x_{t-1})$. The state is partially measured in the measurement space $Z \in R^{nz}$ with likelihood $p(z_t|x_t)$. The posterior density at time t, $p(x_t|z_{1:t})$, which is the probability density of the state x_t at time t given all observations $z_{1:t}=(z_1,\ldots,z_t)$ up to time t containing all the information about the state x_t, can be computed using the Bayes recursion from the probability density $p(x_{t-1}|z_{1:t-1})$ at time t-1.

$$p(x_t|z_{1:t-1}) = \int p(x_t|x_{t-1})p(x_{t-1}|z_{1:t-1})dx_{t-1} \tag{1}$$

$$p(x_t|z_{1:t}) = \frac{p(z_t|x_t)p_{t|t-1}(x_t|z_{1:t-1})}{\int p(z_t|x_t)p_{t|t-1}(x_t|z_{1:t-1})dx_t} \tag{2}$$

In this paper, we assume each target follows a Gaussian model, i.e.

$$p(x_t|x_{t-1}) = N\left(x_t; f_{t|t-1}(x_{t-1}), Q_{t-1}\right) \quad (3)$$

$$p(z_t|x_t) = N\left(z_t; g_t(x_t), R_t\right) \quad (4)$$

where $N(.;m,P)$ denotes a Gaussian density with mean m and covariance P; $f_{t|t-1}(x_{t-1})$, $g_t(x_t)$, Q_{t-1} and R_t are dynamic transition function, likelihood function, process noise covariance and observation noise covariance respectively.

Assume there are m targets moving independently, and described by state vectors at time t: $x_t=\{x_t^1,x_t^2,\ldots,x_t^m\}$. Further assume that target i generates observation $z_t^{[i]}\leftarrow p(z_t^{[i]}|x_t^i)$ and measured with probability P_D. Now, with the measurement vector at time t, $z_t=\{z_t^1,z_t^2,\ldots,z_t^n\}$, where z_t^j is either clutter or one of $\{z_t^{[1]},z_t^{[2]},\ldots,z_t^{[m]}\}$. The task of data association now is to find correspondence between them, and apply above single target tracking filtering process. In practice, because of clutter and noise, on one side, the amount of measurements may be huge, far greater than the number of targets, e.g. $n>>m$. On the other, the measurement of some targets may not be detected, e.g. $P_D<1$. This results in challenges of data association.

3 SDT Method

The core concept of SDT multiple target tracking method is the 'signature' of a target. The intuitive idea of SDT is a simple scenario: on a long fluorescent-lag radar screen, an air craft "seen" by the radar will generate a curve, which is a signature of that air craft. An experienced observer will be able to identify types of that air craft according to it movement and appearance in some case. In that case, signature will be much more than track score and N-scan in MHT. It is a fusion of spectral, spatial and temporal features of targets by using knowledge and models of those targets, just as the experienced observer does when visually identifies air craft types.

The signature of a target here is the numerical representation of characteristics of the corresponding measurement data acquired by the measurement system (or sensor system). The measurement system converts the physical (such as reflectance, radiation, shape, etc) and behavioural (such as attitude, speed, acceleration and turning) characteristics of the target into the numerical data. The low level signature is in the form of the raw measurements. Analysis methods need to be established to extract measures to quantify the signature of targets from various perspectives, and from spectral, spatial and temporal dimension. For example, certain type of aircraft may fly certain speed and acceleration, which can be characterized by Doppler signal. Then the signature fuses the spectral, spatial and temporary features of targets based on the models and knowledge of target properties and dynamics. And the overall belief of a target signature is evaluated based on the Markov property of target dynamics, consistency of consecutive spatial/spectral features, and the detection probability.

Different from the traditional tracking methods, the data association process of SDT occurs before the data filtering process, during which the signature plays a central role. Since the signature fusing spectral, spatial and temporal data, it is the most complete information which can be used to separate targets from clutter and noise. After confirmation of data association by the signature, the input data of state filter, i.e, Kalman filter, contain more signals than noise, thus the accuracy of tracking can be improved greatly. The flow diagram of SDT is given in Fig.1.

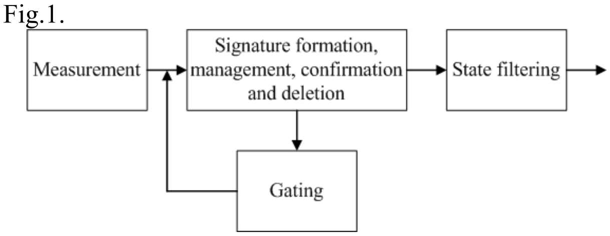

Fig.1 SDT flow diagram

In the rest of this section, we will first define the structure of 'Signature', and then discuss its formation and management, signature fusion and Markov property, and then the tracking algorithm using signature.

3.1 Signature of a target

The signature of a target o is defined as a data structure from time a to time t, where a=max(t-N+1,s), s is the time the target appears, and N is the length of time series to be analyzed:

Signature (o, t) = $\{D_t, L_t, F_t, K_t, C_t, A_t\}$
{
Measurement Sequence: $D_t=(z_a,\ldots,z_t)$;
Pre-State Sequence: $L_t=(\xi_a,\ldots,\xi_t)$;
Feature Sequence: $F_t=(\eta_a,\ldots,\eta_t)$;
Temporal Markov Property: $K_t=\{\zeta_t, \lambda_t\}$;
Feature Markov Property: $C_t=\{\mu_t, \tau_t\}$;
Overall belief of the signature (o): $A_t=\{\theta_t\}$;
};

where t is current time; $z_t=\{z_i^1,z_i^2,\ldots,z_i^M\}$ is data obtained by M sensors; ξ_t is the preliminary estimated state (referred to as pre-state) at time t, consisting of position, velocity and acceleration of the target, and derived from N slices of measurement data D_t either by polynomial fitting or by differentiation. F_t is the spectral and spatial feature trace which is extracted from feature measurement data in the spectral and spatial space according to measurement model. For example, colour histogram can be selected as the feature for video data. The kinematics Markov belief ζ_t represents the temporal consistency of the state at time t to the whole state sequence, while the feature Markov belief μ_t represents the consistency of spectral and spatial feature at time t with the whole feature sequence. λ_t and τ_t are the probabilities that o is regarded as a target based on the respective properties above. It is worth noting that target properties and movement models are used when calculating kinematics Markov belief and feature Markov belief. Creation of target signature is a process of data fusion based on knowledge and models. For example, a car makes turning and moves away from a camera, its picture and location will vary accordingly within N time instances. The overall belief θ_t gives the probability that the signature is generated by a true target.

3.2 Signature-based Gating

The signature of a moving target has Markov property. Between two consecutive sampling time steps, the spectral/spatial features and behavioural characteristics of one unique target are same or similar. For example, in crowd video tracking, since people are moving around, their image features are changing all the time; but this kind of change is continuous, and it satisfies the Markov property. For gating and neighbourhood determination, not only the continuity of motion and movement, but also the continuity of feature will be considered.

The high level feature data derived from original measurement data in spectral and spatial space is denoted as η, while the position data is denoted as z^p. For convenience of discussion the measurement data is denoted as z'=$\{z^p, \eta\}$.

Markov properties of the target location, velocity and acceleration are used to define gating of the signature of a target. If the pre-state of a target at time t-1 is ξ_{t-1}, the kinematics neighbourhood is defined as follows:

$$G_K = \left\{ \begin{array}{l} z' = \{z^P, \eta\}: \\ \left(z^P - g_t\left(f_{t|t-1}(\xi_{t-1})\right)\right)^T S^{-1}\left(z^P - g_t\left(f_{t|t-1}(\xi_{t-1})\right)\right) \le \varepsilon_K \end{array} \right\} \quad (5)$$

where $f_{t1|t-1}(x_t|x_{t-1})$ and $g_t(z_t|x_t)$ are dynamic transition function and likelihood function, respectively. The measurement covariance S depends on f and g, while different choice of dynamic transition function f represents different Markov property. Note here, the gating in SDT is respect to all location, velocity and acceleration, not just location as in other algorithms.

Similarly, we have gating for features:

$$G_F = \left\{z' = \{z^P, \eta\} : c_F\left(\eta|\eta_{t-1}\right) \cdot c_K\left(\eta|F_{t-1}\right) \ge \varepsilon_F \right\} \quad (6)$$

where c_F is the feature similarity function, whose bigger value indicates more similarity of two feature data; c_K is the feature consistency function, whose bigger value indicates more consistency of feature η in feature sequence F_t, this is involved with the trend of feature data from F_t to η.

The overall gating is the intersection of the two:

$$G = G_K \bigcap G_F \quad (7)$$

3.3 Signature formation and management

The signature formation and management here include both tracking and track initialization, which is the most difficult task, and is a necessity for detection and tracking of new targets.

If there is a measurement z_t at time t which does not fall into any neighbourhood (gating) of all existing signatures, a new signature o will be formed with z_t as the first measurement.

At time t+1, the measurement set $\{z^1_{t+1},\ldots,z^{mo}_{t+1}\}$ within the neighbourhood of o is detected after gating. Associating these measurements with o will generate m_o+1 new signatures, o_j, j=0,1,…,m_o, where j=0 indicates the case that at time t+1 o is not detected. Thus one signature extends to m_o+1 new possible signatures. Repeat the same process at time t+2 will result in a tree as shown in Fig. 2.

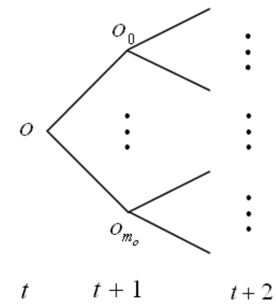

Fig.2 Signature formation.

Thus we maintain a tree, where a path from root to leaf in the tree represents a signature. Among those possible signatures, only one with the overall belief reaching the threshold is considered as true signature of the target.

For the convenience of discussion, the following notations will be used. Θ, the signature tree of time step t-1; $o=\{D_{t-1},L_{t-1},F_{t-1},K_{t-1},C_{t-1},A_{t-1}\}$, one of its leaf node signatures; $Z^o_t=\{z^i_t, i=1,\ldots,m_o\}$, the measurement data set at time t, which falls in the neighbourhood of o; $o_j=\{D_{t,j},L_{t,j},F_{t,j},K_{t,j},M_{t,j},C_{t,j},A_{t,j}\}$, j=0,1,...,m_o, new generated signatures after associating these m_o measurements with o respectively at time step t. Moreover, the measurement set fall in the neighbourhood of tree Θ is

$$Z^\Theta_t = \bigcup_o Z^o_t = \{z^i_t, i = 1,\cdots,m\} \quad (8)$$

3.4 Signature Markov property

The temporal properties of the target signature describe the kinematics of the target within N time instances, which are in terms of location, velocity and acceleration. We believe that the values in N time instances will fit to certain kinematics models, and therefore, we use curve fitting to approximate those kinematics models, and to measure how well the current value fit to the whole models over N time instances. Mathematically, for each signature at t+1, the kinematics Markov belief is given by:

$$\zeta_{t,j} = P\left(\xi_{t,j}|L_{t-1}\right) = c_S\left(\xi_{t,j}|\xi_{t-1}\right) \cdot c_M\left(\xi_{t,j}|L_{t-1}\right) \quad (9)$$

where c_S represents the transition relationship between the pre-state of time t-1 ξ_{t-1} and pre-state of time t $\xi_{t,j}$. The bigger value of c_S denotes more correlation between them. c_M is the function which represents the consistency of pre-state $\xi_{t,j}$ in the whole pre-state sequence L_t. This is involved with the continuity of movement, motion, and coordination of model transition.

Spectral and spatial feature sequence of a target in N time instances also show strong Markov property. A turning face image sequence shall change its visual appearance smoothly. In other words, two consecutive images will show visual feature similarities. Therefore the feature similarity measure can be used as a base to calculate the feature Markov belief, which has the similar form to the temporal Markov belief:

$$\mu_{t,j} = P\left(\eta_{t,j} \middle| F_{t-1}\right) = c_F\left(\eta_{t,j} \middle| \eta_{t-1}\right) \cdot c_K\left(\eta_{t,j} \middle| F_{t-1}\right) \quad (10)$$

where c_F is the feature similarity function, whose bigger value indicates more similarity of two feature data; c_K is the feature consistency function, whose bigger value indicates more consistency of feature η in feature sequence F_t.

Spectral and spatial feature measures of the target can be extracted and represented by feature measure vector. Let us take visual tracking as an example, where targets' data are colour images. Colour histogram is widely used as spectral and spatial feature measure of targets. Let x_i, i=1,…,n, denote pixel locations of targets centered at x_0 (take it as 0 when building model), represent colour distribution with discrete m-bin colour histogram, and use $b(x_i)$ denote the colour bin of the target at x_i. Assume the size of target is normalized which results in kernel radius h=1. Then, the probability q of colour u in the model is:

$$q_u = C\sum_{i=1}^{n} k\left(\left\|\frac{x_0 - x_i}{h}\right\|^2\right)\delta[b(x_i) - u)]$$

where C is the normalization constant. With the feature measure vector q, one can choose *Bhattacharyya* coefficient as similarity measures between two target measurements, or between a target candidate and a target template, which can be used to characterize the feature transition probability. This feature will be used in section 4.2.

3.5 Overall belief with signature confirmation and deletion

Given a signature o, it is necessary to ensure that o is generated by a true target, which requires to evaluate the overall belief A_t={θ_t} of the signature.

Let us work out the criteria for the decision making. Given the previous overall belief of time t-1,

$$\theta_{t-1}=\omega_{L,t-1}\lambda_{t-1}+\omega_{F,t-1}\tau_{t-1},\ \omega_{L\ t-1}+\omega_{F,t-1}=1 \quad (11)$$

Then the overall belief of current time t is derived as followed.

Given the kinematics Markov belief as shown in (9), the kinematics likelihood ratio is:

$$\rho_{t,j} = \begin{cases} \dfrac{P_D}{\beta}\zeta_{t,j}, j \neq 0 \\ \dfrac{1-P_D}{1-\beta}, j = 0 \end{cases} \quad (12)$$

where P_D is the detection probability; β is the clutter density. Through the Sequential Probability Ratio Test (SPRT) rules we have

$$\lambda_{t,j} = \frac{\rho_{t,j} \cdot \lambda_{t-1}}{\rho_{t,j} \cdot \lambda_{t-1} + \left(1-\lambda_{t-1}\right)} \quad (13)$$

Similarly, given feature Markov belief as shown in (10), the feature likelihood ratio is:

$$\gamma_{t,j} = \begin{cases} \dfrac{P_D}{\beta}\mu_{t,j}, j \neq 0 \\ \dfrac{1-P_D}{1-\beta}, j = 0 \end{cases} \quad (14)$$

Thus we have

$$\tau_{t,j} = \frac{\gamma_{t,j} \cdot \tau_{t-1}}{\gamma_{t,j} \cdot \tau_{t-1} + \left(1-\tau_{t-1}\right)} \quad (15)$$

In a summary, the overall belief of signature o_j is given by

$$\theta_{t,j}= \omega_{L,t}\lambda_{t,j}+\omega_{F,t}\tau_{t,j},\ \omega_{L,t}+\omega_{F,t}=1 \quad (16)$$

With signature confirmation and deletion threshold P_{TC} and P_{TD}, the Bayesian confirmation and deletion logic is given below.

$$\begin{cases} \theta_{t,j} < P_{TT}, \quad Delete \\ P_{TT} < \theta_{t,j} < P_{TC}, Further\ Investigation \\ \theta_{t,j} > P_{TC}, \quad Confirm \end{cases} \quad (17)$$

Based on above discussion, the complete SDT tracking algorithm can be summarized as in Tab.1.

4 Experimental Results

Experiments have been conducted using both simulated radar data and visual data. The performance evaluation is conducted using criteria in [7]. To evaluate the performance, SDT tracking results are compared with those of MHT. The MHT algorithm used here was given by Cox and Hingorani [8], and more details can be found in the literature. Performance comparison between SDT and MHT is on the same data sets.

4.1 Tracking of simulated radar data

The radar data of multiple aircrafts are generated by Monte Carlo based computer simulation. The aircraft's state vector includes the location, velocity and acceleration in three dimensions of the inertial Cartesian frame (i.e. the flat earth reference frame). For each trajectory, the initial state vector is randomly generated, and then an entire trajectory is generated, which consists of several alternating maneuvering and non-maneuvering sections. The duration of a maneuver section is uniformly distributed on an interval [10s, 20s], as suggested by Singer [9]. The values of command variables, including height, horizontal speed and heading angle to be achieved at the end of this section, are randomly generated. On the other hand, an exponential distribution is used to model the duration of a non-maneuvering interval with the expectation of the distribution as a user-specified design parameter

In simulations, clutter is modelled as non-fluctuating discrete point scatters uniformly distributed in the surveillance volume with RCS amplitudes varying from 10^2 to 10^6 m^2. The occurrence of clutter measurements is modelled as a Poisson process.

Tab.1 SDT algorithm

Given: tree set $\Xi_t=\{\Theta^{(k)}_t|k=1,\ldots,n_{t-1}\}$ at time t-1, measurement set Z_t at time t, and multiple target state set $X_s=\{x_{s,i}|i=1,\ldots,o_s\}$ at time $s=t-N$, where

$\Theta^{(k)}_{t-1}=\{o^{(k,l)}_t|l=1,\ldots,n^{(k)}_{t-1}\}$,

$o^{(k,l)}_{t-1}=\{D^{(k,l)}_{t-1},L^{(k,l)}_{t-1},F^{(k,l)}_{t-1},K^{(k,l)}_{t-1},C^{(k,l)}_{t-1},A^{(k,l)}_{t-1}\}$.

Step 1.(Prediction)

$\xi^{(k,l)}_{t|t-1}=f_{t|t-1}(\xi^{(k,l)}_{t-1})$; for $l=1,\ldots,n^{(k)}_{t-1}$ and $k=1,\ldots,n_{t-1}$.

Step 2.(Gating)

For $k=1,\ldots,n_{t-1}$, do gating using (7) to get $Z^{(k,l)}_t$, $l=1,\ldots,n^{(k)}_{t-1}$, and derive $Z^{(k)}_t$ for $\Theta^{(k)}_{t-1}$ through (8).

$Y_t=Z_t/(\cup_k Z^{(k)}_t)$.

Step 3.(Signature formation)

Associate $o^{(k,l)}_{t-1}$ with $Z^{(k,l)}_t\cup\Phi$ and form new $1+m^{(k,l)}_t$ signatures $o^{(k,l)}_{t,j}$, where $m^{(k,l)}_t=|Z^{(k,l)}_t|$, and Φ is the empty set.

$o^{(k,l)}_{t,j}=\{D^{(k,l)}_{t,j},L^{(k,l)}_{t,j},F^{(k,l)}_{t,j},K^{((k,l)}_{t,j},C^{(k,l)}_{t,j},A^{(k,l)}_{t,j}\}$;

$j=0,1,\ldots,m^{(k,l)}_t$; $l=1,\ldots,n^{(k)}_{t-1}$; $k=1,\ldots,n_{t-1}$; {

derive $\xi^{(k,l)}_{t,j}$ from $D^{(k,l)}_{t,j}$ by polynomial fitting; derive $\zeta^{(k,l)}_{t,j}$, $\mu^{(k,l)}_{t,j}$, ${}^{(k,l)}_{t,j}$, $\tau^{(k,l)}_{t,j}$, and ${}^{(k,l)}_{t,j}$ respectively through (9), (10) ,(13), (15) and (16);}

If $Y_{t+1}\neq\Phi$ and $Y_{t+1}=\{z^j_t,j=1,\ldots,m_{Y(t)}\}$, initialize $m_{Y(t)}$ signatures $\Theta^{(k)}_{t|t-1}=\{o^{(k,l)}_{t|t-1}|l=1\}$.

Adjust the superscript and subscript.

$\Xi_{t|t-1}=\{\Theta^{(k)}_{t|t-1}|k=1,\ldots,n_{t|t-1}\}$, where $n_{t|t-1}=n_{t-1}+m_{Y(t)}$.

Step 4.(Signature confirmation and termination)

Initialize $\Xi_t=\Phi$; $n_t=0$;

For $k=1,\ldots,n_{t|t-1}$, do signature confirmation, termination and management, and adjust the superscript and subscript.

$\Xi_t=\{\Theta^{(k)}_t|k=1,\ldots,n_t\}$; $\Theta^{(k)}_t=\{o^{(k,l)}_t|l=1,\ldots,n^{(k)}_t\}$.

Step 5.(State filtering)

For $k=1,\ldots,n_{t-1}$, $\Theta^{(k)}_t$ with the forming time $s^{(k)}_t$, {

If $s^{(k)}_t\leq t-N+2$ with $\Theta^{(k)}_t$ been confirmed, find $o^{(k,m)}_t$ in $\Theta^{(k)}_t$ with $m=\arg\max_l(\theta^{(k,l)}_t)$:

if the corresponding track has been initialized, find it in X_s and put it along with $y^{(k,m)}_{a(k,m,t)}$ into the state filter to get a new state $x_{s+1,i}$;

else set $y^{(k,m)}_{a(k,m,t+1)}$ as the position initial value and initialize state as $x_{s+1,i}$;}

$X_{s+1}=\{x_{s+1,i}|i=1,\ldots,o_{s+1}\}$;

Output:

$\Xi_t=\{\Theta^{(k)}_t|k=1,\ldots,n_t\}$;

$\Theta^{(k)}_t=\{\alpha^{(k,l)}_t|l=1,\ldots,n^{(k)}_t\}$;

$X_{s+1}=\{x_{s+1,i}|i=1,\ldots,o_{s+1}\}$;

Radar data is simulated using typical air-surveillance radar with signal to noise ration around 10db. Two different simulations are carried out separately. In the first one, the data contains 10 targets in a 3-dimension space of [30km-60km, 30km-60km, 1-5km]. Target maneuvering parameters are: speed: 100-500m/s; acceleration: 0-9g. The MHT and SDT tracking results are shown in Fig.3 and Fig.4 respectively. From Fig.3-4 it can be seen that SDT has longer true trajectories and shorter false trajectories than MHT.

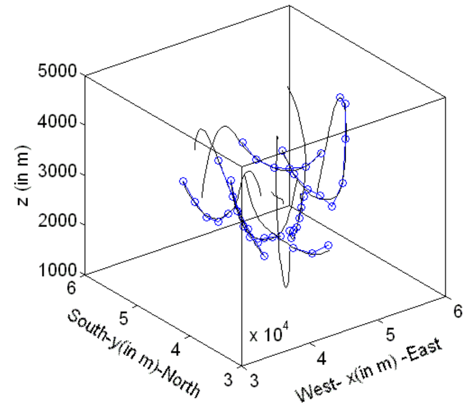

Fig.3 MHT tracking results. The black lines denote the true trajectories of 10 maneuvering aircrafts, while the blue lines with circle markers represent the MHT tracking results

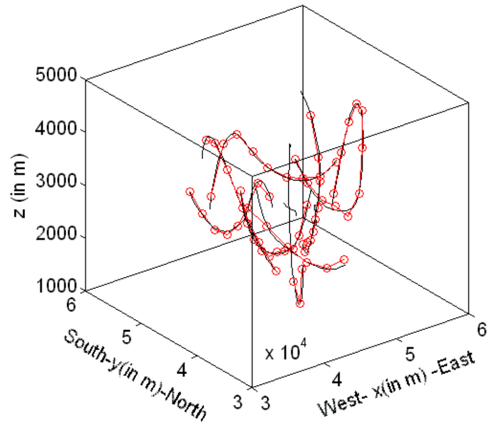

Fig.4 SDT tracking results. The black lines denote the true trajectories of 10 maneuvering aircrafts, while the red lines with circle markers represent the SDT tracking results

In the second simulation, the data contains around 100 targets in a 3-dimension space of [0-100km, 0-100km, 1-8km]. Targets maneuvering parameters are: speed: 0-500m/s, acceleration: 0-10g. A simulated data set is shown in Fig.5. The average performance of SDT and MHT on 100 simulation data is listed in Tab.2.

From the Tab.2, one can see that SDT outperforms MHT in almost all aspects. SDT output more true tracks than MHT (101 vs 82), and less false tracks (0 vs 2). The live time of true tracks of SDT is 14% higher than MHT. The tracking accuracy of SDT is also higher than MHT. This is mainly due to well-defined kinematics properties in the target signature, which are much more sophisticated than track score in MHT. Meanwhile, SDT removes false associations at the earliest time before filtering, which not only significantly reduces the computation, but also improves the capability of the algorithm in processing large amount targets and clutters.

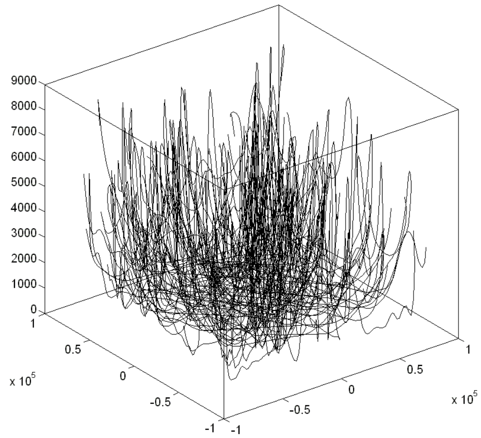

Fig.5. Simulated trajectories of 100 maneuvering aircrafts

Tab.2 Average performance of SDT and MHT

	No. of targets	Tracks detected	True Tracks	False Tracks
MHT	100	84	82	2
SDT	100	101	101	0
	Live Time of True Track (s)	Live Time of False Track (s)	Location RMSE(m)	Speed RMSE(m/s)
MHT	538.9	330	13913	151.4
SDT	617.5	Null	10389	78.06

4.2 Tracking Visual Targets

The visual track experiment was conducted by tracking of very mobile tropical fishes in a fishbowl (90cm×60cm×50cm). To simulate clutters, slips with similar size and colour to fishes were added and waves were created as disturbances. The number of fish varies from 14 to 19; the number of spurious targets (slips) is between 60 and 71. In order to distinguish between fishes and slips, colour histogram together with edge features are used as visual features.

We choose 100 frames to manually identify individual targets and true tracks for both SDT and MHT. The statistics are in the Tab.3. As we can see from the table that the overall successful association and tracking process by SDT is 94.76%, which is much higher than 88.12% by MHT. Fig.6 and Fig.7 show tracking results with trajectories in the form of solid lines of MHT and SDT. It can be seen from the two figures that SDT tracking results have less false positives than those of MHT, and it takes less slips as fishes than MHT. The SDT algorithm also tracks a fish which is in the left bottom of the fishbowl and missed by the MHT method. Moreover, the lives of true tracks of SDT are longer than those of MHT, which also can be seen in Tab.3.

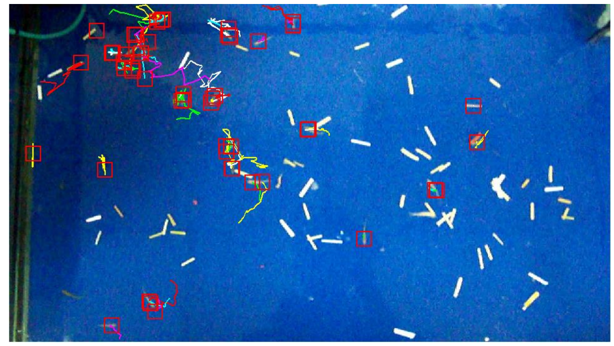

Fig 6. Trajectories of tropical fishes by MHT algorithm

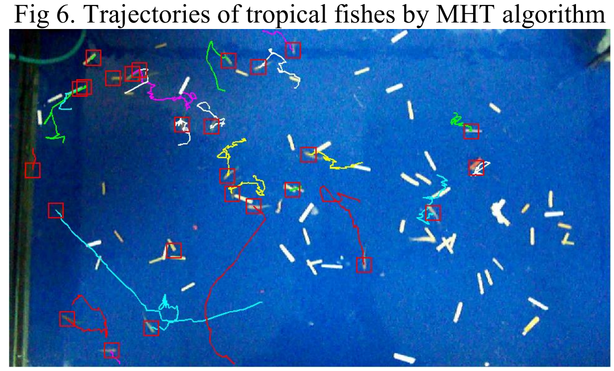

Fig 7. Trajectories of tropical fishes by SDT algorithm

Tab.3 Visual target tracking by SDT and MHT

Target index	No frames target appears	SDT		MHT	
		Track frames	Prob. of true track	Track frames	Prob. of true track
1	22	22	100.00%	15	68.18%
2	92	78	84.78%	73	79.35%
3	100	84	84.00%	80	80.00%
4	100	100	100.00%	97	97.00%
5	86	84	97.67%	82	95.35%
6	100	100	100.00%	92	92.00%
7	100	100	100.00%	92	92.00%
8	100	100	100.00%	99	99.00%
9	100	98	98.00%	95	95.00%
10	100	98	98.00%	77	77.00%
11	100	100	100.00%	92	92.00%
12	100	100	100.00%	95	95.00%
13	100	100	100.00%	91	91.00%
14	100	88	88.00%	80	80.00%
15	64	30	46.88%	29	45.31%
16	59	59	100.00%	50	84.75%
17	83	83	100.00%	83	100.00%
18	39	39	100.00%	38	97.44%
19	21	21	100.00%	20	95.24%
Overall	1566	1484	94.76%	1380	88.12%

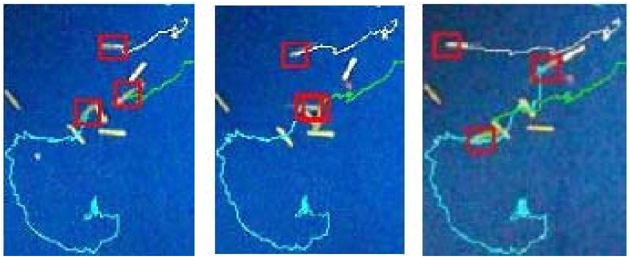

Fig.8. Tracking of crossing targets by SDT

In dealing with targets crossing, SDT method has shown its advantage by fusion of spectral, spatial and temporal information. In Fig.8, the picture on the left shows two fishes moving towards each. The picture in the middle is the merging of these two fishes, while the picture on the right shows that they are travelling on their own directions and SDT estimated their trajectories accurately which reflect true motion of these two fishes.

5 Conclusion

We have presented a novel signature-driven multiple target tracking algorithm, SDT. It makes use of the target signatures at the earliest possible time in the tracking algorithm to remove spurious association. The experimental results have shown its superior performance. We have noticed that fact that SDT has higher tracking accuracy than MHT. This may due to the fact that in SDT, filtering is performed after confirmation and deletion of signature. The filtering results are not affected by spurious targets data.

Further work will be on the fusion of multiple sensory information, which may be heterogeneous and at different locations.

Acknowledgment

The authors wish to acknowledge the support by the members of Sensor Networks and Applications Research Center (SNARC). This work is partially supported by the National Natural Science Foundation of China (Grant No. 60772154).

References

[1] Kirubarajan, T., Bar-shalom, Y, *Probabilistic data association techniques for target tracking in clutter*, Proceedings of the IEEE, Vol 92, No. 3, pp. 536–557, Mar 2004.

[2] Donald B. Reid, *An Algorithm for Tracking Multiple Targets*, IEEE transactions on automatic control, Vol 24, No.6, pp. 843–854, Dec 1979.

[3] S.S. Blackman, *Multiple hypothesis tracking for multiple target tracking*, IEEE Aerospace and Electronic Systems Magazine, Vol 19, No. 1, pp.5-18, Jan 2004.

[4] J. Lancaster and S. Blackman, *Joint IMM/MHT Tracking and Identification for Multi-Sensor Ground Target Tracking*, 2006 9th International Conference on Information Fusion, Florence, Italy, 10-13 July 2006, Proc.2006, pp. 1–7.

[5] H. Jiang, S. Fels, and J. J. Little, *A linear programming approach for multiple object tracking*, Computer Vision and Pattern Recognition 2007, Minneapolis, MN, 17-22 June 2007, Proc.2007, pp. 1-8.

[6] Yuan Li, Chang Huang and Ram Nevatia, *Learning to Associate: HybridBoosted Multi-Target Tracker for Crowded Scene*, Computer Vision and Pattern Recognition 2009, Miami, FL, 20-25 June 2009, Proc.2009, pp. 2953 – 2960.

[7] H. Leung, Zhijian Hu, and M. Blanchette, *Evaluation of multiple radar target trackers in stressful environments*, IEEE. TRANS. AES, Vol 35, No. 2, pp. 663–674, Apr 1999.

[8] I. J. Cox and S. L, *An efficient implementation of Reid's multiple hypothesis tracking algorithm and its evaluation for the purpose of visual tracking*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 18, No. 2, pp. 138–150, Feb 1996.

[9] Li, X. R., and Jilkov, V. P, *Survey of Maneuvering Target Tracking. Part I: Dynamic Models*, IEEE Transactions on Aerospace and Electronic Systems, Vol.39, No. 4, pp. 1333–1364, Oct 2003.