Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Signature-Driven Multiple Visual Target Tracking

Shuyan Sun, Zhipei Huang, Sheng Jiang, Jiankang Wu, Zhiqiang Zhang
School of Information Science and Engineering
Graduate University of Chinese Academy of Sciences
Beijing, China
jkwu@gucas.ac.cn, sunshuyan07@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 uses target signature in spectral, spatial and temporary spaces as well as the Markov property of target movement, so that the data association process in SDT is very efficient and effective. The experimental results have shown outstanding performance.**

***Keywords*—tracking, filtering, estimation, signature, sensor fusion**

I. INTRODUCTION

Data association for multiple target tracking has always been a challenging research topic. Multiple Hypothesis Tracking (MHT)[1,2] is regarded as the best multiple target tracking algorithm. The fundamental idea is to make use of the temporal information of the target measurement through the track scores which are accumulated over time. However, it is based on hypothesis-test principle. The performance becomes poor when the number of targets and amount of clutter increase-too many hypotheses complicate the algorithm.

We have proposed and developed a Signature-Driven multiple target Tracking (SDT) algorithm which uses the spectral, spatial and temporary features of the target to reduce the number of hypotheses and to increase the adaptability of the algorithm.

In what follows, section II provides the background for multiple target tracking, while section III describes our SDT tracking algorithm. The experimental results are presented in section IV. Finally, section V is the conclusion.

II. BACKGROUND

The multiple target tracking is based on single target tracking together with data association for multiple measurements.

A. Single Target Tracking

In single target tracking, the state of a target is assumed to be a first order Markov process on the state space $\chi \in R^{nx}$ with transition density $p(x_t|x_{t-1})$[3]. This Markov process 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.

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

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

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

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

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

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.

B. Modeling of image and video sequences

In the area of image and video processing[4,5], modeling and analysis of image and video sequences has been a practical solution for the extraction and tracking of targets. Before launching the tracking algorithm, a target model base should be properly defined and established. This may involve the study of sensory system as well as the characteristics of target sensory data with consideration of reflectance, target dynamics and the characteristics of the clutter, etc. After that, a model base is to be established. This model base should include the information about the targets and clutter, such as, types of targets, the signature, possible motion models, probability of occurrence of each type of targets, the types of clutter, the signature of those clutter and the probability of occurrence. In order to derive the target model parameters, we may need to collect training data samples and design proper training algorithms. For example, certain type of aircraft may fly at certain speed and acceleration, which can be characterized by Doppler signal.

In this paper, visual tracking of multiple fish among many spurious targets in similar color and shape is taken as an example. The detail will be given below.

III. SDT ALGORITHM

SDT multiple target tracking algorithm starts from the 'Signature' of a target. In this section, we will first define the term 'Signature', and then will discuss the method of the extraction of temporal, spatial and spectral features and other aspects of the SDT algorithm.

A. *Signature of a target*

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

Signature (o, t)=$\{L_t,F_t,K_t,C_t,A_t\}$={
 Features at time t {
  Measurement Sequence: $L_t=(y_a,\ldots,y_t)$;
  Feature Sequence: $F_t=(\eta_a,\ldots,\eta_t)$;
  Temporal Signature: $K_t=\{\xi_t,\zeta_t,\lambda_t\}$;
  Spectral/spatial Signature: $C_t=\{\mu_t,\tau_t\}$;
 Overall belief of the signature (o) at time t: $A_t=\{\theta_t,\nu_t\}$;
  $\theta_t=\omega_{L,t}p(\xi_t|L_t)+\omega_{F,t}p(\eta_t|F_t)$,
  $\nu_t=\omega_{L,t}P(T|L_t)+\omega_{F,t}P(T|F_t)$,
  $\omega_{L,t}+\omega_{F,t}=1$
 };

where j(j=a:t) is the index of time, and t is current time. The equivalent status ξ_t at time t, which contains the information about the target's position, velocity and acceleration, can be derived from N slices of measurement data L_t either by polynomial fitting or by difference, depending on the magnitude of the measurement noise. The kinetics Markov belief $\zeta_t=p(\xi_t|L_t)$ and the feature Markov belief $\mu_t=p(\eta_t|F_t)$ are used to represent the temporal property and the spectral/spatial property of the target. $\lambda_t=P(T|L_t)$ and $\tau_t=P(T|F_t)$ are the probabilities that o is regarded as a target based on the respective properties above.

B. *Signature Gating*

Markov properties of the target position, velocity and acceleration are used to define gating of the signature of a target. If the equivalent status of the signature at time t is ξ_t, the neighborhood of this target is defined as follows:

$$G=\left\{z \mid \left(z-g_t\left(f_{t+1|t}\left(\xi_t\right)\right)\right)^T S^{-1}\left(z-g_t\left(f_{t+1|t}\left(\xi_t\right)\right)\right)\le\varepsilon\right\} \qquad (5)$$

where $f_{t+1|t}(x_{t+1}|x_t)$ and $g_t(z_t|x_t)$ are respectively dynamic transition function and likelihood function, along with measurement covariance S depending on f and g, while different choice of dynamic transition function f represents different Markov property.

C. *Signature formation and management*

If there exists a position measurement z_t at time t which does not fall into any neighborhood of all existing signatures, a new signature α will be formed with z_t as the first measurement of α.

At time t+1, the measurement set $\{z^1_{t+1},\ldots,z^{m\alpha}_{t+1}\}$ within the neighborhood of α will be settled after gating. Associating these measurements with α, new m_α+1 signatures, α_j, j=0,1,...,m_α, where j=0 indicates that all of the measurements are clutter, will be formed and α will be deleted. Thus one signature extends to m_α+1 new signatures and all of them represent only one target with the same emergence time t. Repeat the same implementation at time t+2.

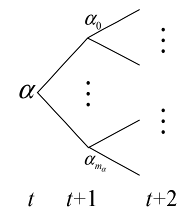

Figure 1. Signature formation.

Thus we maintain a tree consisting of signatures. Each node of this tree denotes only one signature. Each branch from root to leaf node represents time-slices of measurement and all the branches of this tree correspond to only one target. Each signature will be confirmed if it is generated from true target measurement or will be terminated if it is composed of clutter (see section III F). Only one branch of each tree will be chosen as the output of this tree. A tree is confirmed as long as any signature in it has been confirmed. If one confirmed tree has been existed for consecutive N time steps, the confirmed leaf signature with the highest overall belief θ will be sent to the state filter as the output of this confirmed tree.

D. *Temporal signature*

Given a tree structure Θ at time t, $\alpha=\{L_t,F_t,K_t,C_t,A_t\}$ belongs to Θ. At time t+1, with the measurement set $Z^{\alpha}_{t+1}=\{z^i_{t+1},i=1,\ldots,m_\alpha\}$ in the neighborhood of α, the measurement set that falls into the neighborhood of Θ is:

$$Z^{\Theta}_{t+1}=\bigcup_{\alpha}Z^{\alpha}_{t+1}=\left\{z^i_{t+1},i=1,\cdots,m\right\}. \qquad (6)$$

After associated these measurements in Z^{α}_{t+1} with α and formed new m_α+1 signatures, $\alpha_j=\{L_{t,j},F_{t,j},K_{t,j},C_{t,j},A_{t,j}\}$, j=0,1,...,m_α, the kinetics Markov belief of each new signature which represents temporal property is given by

$$P\left(\xi_{t+1,j}\left|Z_{1:t+1}\right.\right)=\begin{cases}\dfrac{1}{c}\cdot P_D\cdot P\left(\xi_{t+1,j}\left|\xi_t\right.\right)\cdot P\left(\xi_t\left|Z_{1:t}\right.\right), & j\neq 0\\ \dfrac{1}{c}\cdot\beta\cdot\left(1-P_D\right)\cdot P\left(\xi_t\left|Z_{1:t}\right.\right) & ,j=0\end{cases} \qquad (7)$$

where P_D is the detection probability; β is the clutter density; and the denominator c is derived by summing over all the numerators whose total number is

$$\sum_{\alpha\in\Theta}\left(1+m_\alpha\right). \qquad (8)$$

As for the temporal information, only N time slides of data up to the current time t will be considered, thus we have

$$\zeta_{t+1,j}=P\left(\xi_{t+1,j}\left|L_{t+1,j}\right.\right)=P\left(\xi_{t+1,j}\left|Z_{1:t+1}\right.\right) \qquad (9)$$

$$P\left(\xi_{t+1,j}\left|\xi_t\right.\right)=N\left(z^j_{t+1}-g_t\left(f_{t+1|t}\left(\xi_t\right)\right),S\right) \qquad (10)$$

where $f_{t+1|t}(x_{t+1}|x_t)$, $g_t(z_t|x_t)$ and S are the same as above.

E. *Spectral/spatial signature*

In this paper, visual tracking of multiple fish among many spurious targets in similar color and shape is taken as an example. Color histogram is used as invariant feature with respect to scale and perspectives. This can be used as the spectral and spatial signature of the target defined in [5].

Let x_i, i=1,…,n, denote pixel locations of model centered at x (take it as 0 when building model). Represent color distribution with discrete m-bin color histogram. Let $b(x_i)$ denote the color bin of the color at x_i. Assume the size of model is normalized which results in kernel radius h=1. Then, the probability q of color u in the model is:

$$q_u = C\sum_{i=1}^{n} k(\| \frac{x_0 - x_i}{h} \|^2)\delta[b(x_i - u)] \tag{11}$$

where C is the normalization constant:

$$C = \left[\sum_{i=1}^{n} k\left(\left\| \frac{x - x_i}{h} \right\|^2 \right) \right]^{-1} \tag{12}$$

while q_u is taken here as the spectral/spatial signature μ_t.

Then ISODATA[7] algorithm is used here to cluster the spectral and spatial signature of samples to distinguish the models (templates) of fish and the spurious targets. Scan each video frame with pyramid structure. For each possible target location, calculate the color histogram and perform template matching using *Bhattacharyya* coefficient which is defined as:

$$\rho(p_u, q_u) = \sum_{u=1}^{m} \sqrt{p_u q_u} \,. \tag{13}$$

Meanwhile, the local maximum of the *Bhattacharyya* coefficient is also taken as the likelihood of the target.

$$P\left(\eta_{t+1,j} \middle| \eta_t\right) = \sum_{u=1}^{m} \sqrt{\eta_{t+1,j}^{(u)} \cdot \eta_t^{(u)}} \tag{14}$$

Then we obtain (15) in the same way as (7)

$$\mu_{t+1,j} = P\left(\eta_{t+1,j} \middle| F_{t+1,j}\right) = \begin{cases} \frac{1}{c} \cdot P_D \cdot P\left(\eta_{t+1,j} \middle| \eta_t\right) \cdot P\left(\eta_t \middle| F_t\right), j \neq 0 \\ \frac{1}{c} \cdot \beta \cdot (1 - P_D) \cdot P\left(\eta_t \middle| F_t\right), j = 0 \end{cases} \tag{15}$$

F. Bayesian confirmation and termination

Given a signature α, it is necessary to ensure that α is generated by a true target[6], which requires the probability $\lambda_t = P(T|L_t)$ and $\tau_t = P(T|F_t)$. Through Bayes Rule,

$$P(T/\alpha) = \frac{P(\alpha/T)P_0(T)}{P(\alpha)} = \frac{P(\alpha/T)P_0(T)}{P(\alpha/T)P_0(T) + P(\alpha/F)P_0(F)} \tag{16}$$

$$P_0(T) + P_0(F) = 1 \tag{17}$$

where $P(T|\alpha)$ and $P(F|\alpha)$ are respectively the probability that α generated by a true target or by clutter. Let $\rho_t = P(\alpha|T)/P(\alpha|F)$ to be the likelihood ratio, then

$$P(T/\alpha) = \frac{\rho_t \cdot P_0(T)}{\rho_t \cdot P_0(T) + (1 - P_0(T))} \tag{18}$$

Given $P(T|\alpha)$ at time t, α will extend to $m_\alpha+1$ new signatures α_j, j=0,1,…,m_α, at time t+1, with the corresponding likelihood ratio $\rho_{t+1,j}$, then we get $P(T|\alpha_j)$

$$P(T/\alpha_j) = \frac{\rho_{t+1,j} \cdot P(T/\alpha)}{\rho_{t+1,j} \cdot P(T/\alpha) + (1 - P(T/\alpha))} \tag{19}$$

With confirmation and termination threshold P_{TC} and P_{TT}, the Bayesian confirmation and termination logic is given below.

$$\begin{cases} P(T/\alpha_j) < P_{TT} & \text{,Signature terminated} \\ P_{TT} < P(T/\alpha_j) < P_{TC}, & \text{continue test} \\ P(T/\alpha_j) > P_{TC} & \text{, Signature confirmed} \end{cases} . \tag{20}$$

Given α and $P(T|\alpha)=\omega_{L,t}P(T|L_t)+\omega_{F,t}P(T|F_t)$ at time t, where $\alpha=\{L_t,F_t,K_t,C_t,A_t\}$ and $\omega_{L,t}+\omega_{F,t}=1$. At time t+1, α extends to α_j, j=0,1,… ,m_α, with corresponding ratio $\rho_{t+1,j}$ and $\gamma_{t+1,j}$

$$\rho_{t+1,j} = \begin{cases} \frac{P_D}{\beta} P\left(\xi_{t+1,j} \middle| \xi_t\right) \\ \frac{1 - P_D}{1 - \beta} \end{cases} \tag{21}$$

$$\gamma_{t+1,j} = \begin{cases} \frac{P_D}{\beta} P\left(\eta_{t+1,j} \middle| \eta_t\right) \\ \frac{1 - P_D}{1 - \beta} \end{cases} \tag{22}$$

then we have

$$P(T/L_{t+1,j}) = \frac{\rho_{t+1,j} \cdot P(T/L_t)}{\rho_{t+1,j} \cdot P(T/L_t) + (1 - P(T/L_t))} \tag{23}$$

$$P(T/F_{t+1,j}) = \frac{\gamma_{t+1,j} \cdot P(T/F_t)}{\gamma_{t+1,j} \cdot P(T/F_t) + (1 - P(T/F_t))} . \tag{24}$$

G. Overall belief of the Signature

Given temporal signature (7), (23) and spectral/spatial signature (15), (24), the overall belief of a Signature is derived by

$$\theta_t = \omega_{L,t} p(\xi_t|L_t) + \omega_{F,t} p(\eta_t|F_t) \tag{25}$$

$$\nu_t = \omega_{L,t} P(T|L_t) + \omega_{F,t} P(T|F_t) \tag{26}$$

where $\omega_{L,t}$ and $\omega_{F,t}$ are weights on the condition $\omega_{L,t}+\omega_{F,t}=1$.

IV. EXPERIMENTAL RESULTS

In our experiment, fish are tracked in a fishbowl (90cm×60cm×50cm). The data are gathered by a video camera with the resolution 1280×720 pixels. There are two kinds of tropical fish whose colors are either yellow or brown, and their length is about 2-3 cm. To complicate the conditions of detection, slips are used here as spurious targets/clutter and the water waves are created as disturbances.

Because the color of slips is very similar to that of fish, color histogram alone as the invariant feature with respect to scale and perspectives in the experiment is found inadequate. So edge feature is added here to distinguish fish from slips.

In this experiment, we assume each target follows a linear Gaussian dynamical model, i.e.

$$p(x_t | x_{t-1}) = N(x_t; F_{t-1}x_{t-1}, Q_{t-1}) \tag{27}$$

$$p(z_t | x_t) = N(z_t; H_t x_t, R_t) \tag{28}$$

TABLE I. THE COMPLETE ALGORITHM

Given: tree set $\Xi_t=\{\Theta^{(k)}_t|k=1,\ldots,n_t\}$ at time t, measurement set Z_{t+1} at time $t+1$, and multiple target state set $X_s=\{x_{s,i}|i=1,\ldots,o_s\}$ at time $s=t\text{-}N+1$, where $\Theta^{(k)}_t=\{\alpha^{(k,l)}_t|l=1,\ldots,n^{(k)}_t\}$, $\alpha^{(k,l)}_t=\{L^{(k,l)}_t, F^{(k,l)}_t, K^{(k,l)}_t, C^{(k,l)}_t, A^{(k,l)}_t\}$.

Step 1.(Prediction)

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

Step 2.(Gating)

For $k=1,\ldots,n_t$, do gating using (5) to get $Z^{(k,l)}_{t+1}$, $l=1,\ldots,n^{(k)}_t$, and derive $Z^{(k)}_{t+1}$ for $\Theta^{(k)}_t$ through (6).

$Y_{t+1}=Z_{t+1}/(\cup_k Z^{(k)}_{t+1})$.

Step 3.(Signature formation)

Associate $\alpha^{(k,l)}_t$ with $Z^{(k,l)}_{t+1}\cup\Phi$ and form new $1+m^{(k,l)}_{t+1}$ signatures $\alpha^{(k,l)}_{t+1,j}$, where $m^{(k,l)}_{t+1}=|Z^{(k,l)}_{t+1}|$. $\alpha^{(k,l)}_{t+1,j}=\{L^{(k,l)}_{t+1,j}, F^{(k,l)}_{t+1,j}, K^{(k,l)}_{t+1,j}, C^{(k,l)}_{t+1,j}, A^{(k,l)}_{t+1,j}\}$; $j=0,1,\ldots,m^{(k,l)}_{t+1}$; $l=1,\ldots,n^{(k)}_t$; $k=1,\ldots,n_t$; {

derive $\xi^{(k,l)}_{t+1,j}$ from $L^{(k,l)}_{t+1,j}$ by polynomial fitting; derive $\zeta^{(k,l)}_{t+1,j}$, $\lambda^{(k,l)}_{t+1,j}$, $\mu^{(k,l)}_{t+1,j}$ and $\tau^{(k,l)}_{t+1,j}$ respectively through (7), (15) ,(23) and (24);}

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

Adjust the superscript and subscript.

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

Step 4.(Signature confirmation and termination)

Initialize $\Xi_{t+1}=\Phi$; $n_{t+1}=0$;

For $k=1,\ldots,n_{t+1|t}$, do signature confirmation, termination and management through (20) and III.C, F, and adjust the superscript and subscript. $\Xi_{t+1}=\{\Theta^{(k)}_{t+1}|k=1,\ldots,n_{t+1}\}$; $\Theta^{(k)}_{t+1}=\{\alpha^{(k,l)}_{t+1}|l=1,\ldots,n^{(k)}_{t+1}\}$.

Step 5.(State filtering)

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

If $s^{(k)}_{t+1}\leq t\text{-}N+2$ with $\Theta^{(k)}_{t+1}$ been confirmed, find $\alpha^{(k,m)}_{t+1}$ in $\Theta^{(k)}_{t+1}$ with $m=\arg\max_l(\theta^{(k,l)}_{t+1})$:

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

$\Xi_{t+1}=\{\Theta^{(k)}_{t+1}|k=1,\ldots,n_{t+1}\}$;

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

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

where $N(.;m,P)$ denotes a Gaussian density with mean m and covariance P, and F_{t-1} and H_t are respectively dynamic transition matrix and observation matrix. The state $x_t=[p_{x,t},p_{y,t},v_{x,t},v_{y,t}]^T$ of each target consists of position $(p_{x,t},p_{y,t})$ and velocity $(v_{x,t},v_{y,t})$, while the measurement is the position with noise. The dynamic transition matrix F_{t-1} and process noise covariance Q_{t-1} are

$$F_{t-1}=\begin{bmatrix} I_2 & \Delta I_2 \\ 0_2 & I_2 \end{bmatrix}, \quad Q_t=\sigma_w^2\begin{bmatrix} \frac{\Delta^4}{4}I_2 & \frac{\Delta^3}{2}I_2 \\ \frac{\Delta^3}{2}I_2 & \Delta^2 I_2 \end{bmatrix} \tag{29}$$

where I_2 and 0_2 denote 2×2 identity and zero matrices, $\triangle$ =1/30s is the sampling interval, and σ_w=5000(pixel/s^2) is the standard deviation of the process noise. The observation matrix H_t and measurement noise covariance matrix R_t are given by $H_t=[I_2,0_2]$, $R_t=\sigma_v^2 I_2$, where σ_v=20(pixel) is the standard deviation of the measurement noise. The detection probability of each target P_D is 0.99, and the clutter density $\beta=2\times10^{-8}$(pixel)$^{-2}$.

Figure 2. Deteced result after distinguishing fish from clutter.

Figure 3. Tracking result without trajectory.

Fig.2 shows the detected result after distinguishing fish from clutter, i.e., slips, utilizing the difference between the spectral/spatial signature of fish and that of slips. In the picture, the red circles indicate that there may be fish. However, clearly there are 9 slips miss-detected as fish in this frame. Fig.3 shows tracking result without trajectory where red squares represent the position estimate of fish. As we can see, by adding temporary feature, miss-detections are erased.

Figure 4. Tracking result with trajectory.

Fig.4 shows tracking results with trajectory in the form of solid lines. The number of fish varies from 14 to 18; the number of spurious targets (slips) is between 60 and 71, which can be seen in Fig.5. The average track loss is 3.6%.

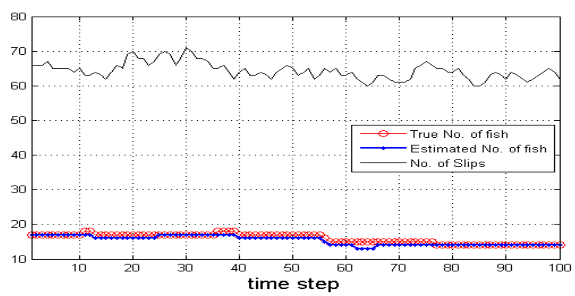

Figure 5. True No.(red circle line), estimated No.(blue point line) of fish and No.of slips(black solid line)

While dealing with crossing targets, our method has proven its advantage by employing fusion of temporal, spectral and spatial information, which has been shown in Fig.6. The picture on the left denotes two targets moving towards each other whose trajectories are painted respectively in light blue and green. The picture in the middle represents the merging of these two fish, while the picture on the right shows that they are traveling on their own directions and our estimated trajectories accurately reflect their motion.

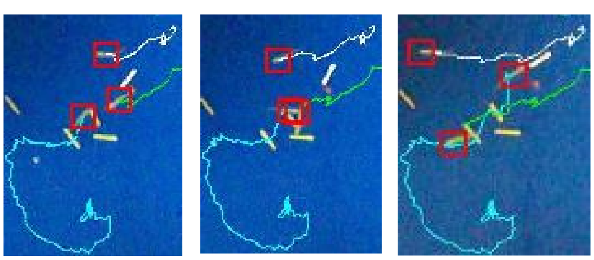

Figure 6. Example of crossing targets.

From the position estimate in Fig.3, the trajectory estimate in Fig.4 and the No. estimate in Fig.5, the SDT multiple target tracking method has shown accurate tracking performance. Making use of target signature in spectral, spatial and temporary spaces as well as the Markov property of target movement from the raw data at earlier time, the filter was able to distinguish targets from clutter as early as possible, so that the data association process in SDT is very efficient and effective.

V. CONCLUSION

We have presented a new multiple target tracking algorithm. It makes use of the target signatures in the earliest possible time, and has shown the superior performance. Further work will be on the testing and improvement of the algorithm.

ACKNOWLEDGMENT

The authors 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) and by the President Foundation of Graduate University of Chinese Academy of Sciences (Grant No. 085102GN00).

REFERENCES

[1] S.S. Blackman, "Multiple hypothesis tracking for multiple target tracking," IEEE Aerospace and Electronic Systems Magazine, Vol. 19(1), pp. 5-18, January 2004.

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

[3] Vo Ba-Ngu and Ma Wing-Kin. "The Gaussian Mixture Probability Hypothesis Density Filter,"IEEE Transactions. Signal Processing, Vol.54(11), pp.4091-4104, November 2006.

[4] Y. Wang, K.-F. Loe, T. Tan, and J.-K. Wu, "Spatio-temporal video segmentation based on graphical models," IEEE Trans. Image Processing, Vol.14(7), pp. 937- 947, July 2005.

[5] D. Comaniciu, V. Ramesh, and P. Meer., "Real-time tracking of non-rigid objects using mean shift," IEEE Proc. on Computer Vision and Pattern Recognition,Vol. 2, pp. 142 – 149, June 2000.

[6] Jing Zhongliang, Zhou Hongren, Wang Peide, "Tracking initiation and termination of multiple maneuvering targets in a dense multi-return environment,"IEEE Conference on Decision and Control, Vol. 4, pp. 2270-2275, December 1990.

[7] Roger del Moral, "Vegetation clustering by means of isodata: Revision by multiple discriminant analysis."Plant Ecology, Springer Netherlands, vol. 29(3), pp. 179-190, January 1975.