Technical Report 2008-003

A Sequential Monte-Carlo and DSmT Based Approach for Conflict Handling in case of Multiple target Tracking

Yi Sun, Layachi Bentabet
Department of Computer Science
Bishop's University
(ysun053, lbentabe@ubishops.ca)

Abstract. In this paper, we propose an efficient and robust multiple target tracking method based on particles filtering and Dezert-Smarandache theory. A model of cue combination is designed with plausible and paradoxical reasoning. The proposed model can resolve the conflict and paradoxes that arise between measured cues due to the partial or total occlusion. Experimental results demonstrate the efficiency and accuracy of the model in case of tracking with multiple cue.

Key Words. Multiple target Tracking, Sequential Monte-Carlo, Dezert-Smarandache Theory, Cue Combination.

(c) Department of Computer Science, Bishop's University
http://cs.ubishops.ca/research.html

1 Introduction

Visual tracking of moving targets has become one of the primary research issues in computer science, due to the increasing demand for reliable activity monitoring and surveillance systems. In order to achieve accurate tracking, several cues have been explored in the literature. These cues include color histogram, edges, motion, camera geometry (field of view), or velocity. As pointed out by [1], individual cue can potentially fail or provide paradoxical interpretations due to the occlusion problem in cluttered scenes and the changes in the illumination which are unavoidable in the real world. Numerous methods suggest the use of multiple cue to increase visual robustness in cases of complex scenes. The integration of the extracted cues into an object representation has been performed using probabilistic methods [1][2] as well as non-probabilistic methods [3]. With a knowledge of a priori distributions and conditional probabilities, the probabilistic methods, and especially the Bayesian inference, offer the most complete, scalable and theoretically justifiable approach for data fusion. However, in real complex scenes such complete knowledge is difficult to obtain due to high occlusion, background clutter, illumination and camera calibration problems. In [4], an alternative approach is proposed for data fusion; namely the Dempster-Shafer (DST) theory. DST theory does not require the knowledge of prior probabilities. The uncertainty and imprecision of a body of knowledge are represented via the notion of confidence values that are committed to a single or a union of hypotheses. The orthogonal sum rule of DST theory allows the integration of information from different sources into a single and overall representation. Unfortunately, Bayesian inference and Dempster-Shafer theories lack to provide an interesting manner of modeling conflicts and paradoxical interpretation arising between the difference information sources. The Bayesian inference assumes that all sources provide bodies of evidences using the same objective and universal interpretation of the phenomena under consideration; therefore, it cannot handle conflicts [5]. In most practical fusion applications based on DST theory, ad-hoc or heuristic techniques must always be added to the fusion process to manage or reduce the possibility of high degree conflict between the sources. Otherwise, the fusion leads to inaccuracies or cannot provide a reliable result at all. To overcome these limitations, a recent theory of plausible and paradoxical reasoning has been developed in [5]. The Dezert-Smarandache Theory (DSmT) can be considered as a generalization of the DST theory. In this paper, we propose a sequential particles filter approach for multiple target tracking using multiple cue. The different cues are combined using the DSmT theory. If the targets are partially or completely occluded, the conflicts and paradoxes that arise between the measured cues are assessed and used in the tracking process. The proposed scheme is simple and provides an effective tracking in cluttered scenes.

2 The Dezert-Smarandache Theory

The Dezert-Smarandache Theory (DSmT) of plausible and paradoxical reasoning [5] is a generalization of the classical Dempster-Shafer theory (DST), which allows formal combining of rational, uncertain and paradoxical sources. The DSmT is able to solve complex fusion problems where the DST usually fails, especially when conflicts between sources become large. In this section, we will first review the principle of the DST before discussing the fundamental aspects of the DSmT.

2.1 Principle of Dempster-Shafer Theory

The DST makes inferences from incomplete and uncertain knowledge by combining additional sources of confidence, even in the process of partially contradictory sensors. The DST contains the Bayesian theory of partial belief as a special case. In the DST, there is a fixed set of mutually exclusive and exhaustive elements, called the frame of discernment, which is symbolized by $\Theta = \{\theta_1, \theta_2, \ldots, \theta_N\}$. The frame of discernment Θ defines the propositions for which the sources can provide confidence. Information sources can distribute mass values on subsets of the frame of discernment, $A_i \in 2^{\Theta}$. If an information source can not distinguish between two propositions A_i and A_j, it assigns a mass value to the set including both hypotheses ($A_i \cup A_j$). The mass distribution for all hypotheses has to fulfill the following conditions

(c) Department of Computer Science, Bishop's University
http://cs.ubishops.ca/research.html

$$\begin{aligned} &0 \le m(A_i) \le 1 \\ &m(\phi) = 0 \\ &\sum_{A_i \in 2^{\Theta}} m(A_i) = 1 \end{aligned} \qquad (\mathbf{1})$$

Mass distributions m_1, m_2,…, m_d from d different sources are combined with Dempster's orthogonal rule. The result is new distribution, $m = m_1 \oplus m_2 \oplus \ldots \oplus m_d$, which carries the joint information provided by the d sources.

$$m(A) = (1-K)^{-1} \sum_{A_1 \cap \ldots \cap A_p = A} \left[\prod_{i=1}^{d} m_i(A_i) \right] \qquad (2)$$

where

$$K = \sum_{A_1 \cap \ldots \cap A_p = \phi} \left[\prod_{i=1}^{d} m_i(A_i) \right].$$

K is a measure of conflict between the sources and is introduced as a normalization factor. The larger *K* is, the more the sources are conflicting and the less the combination has sense. Two functions can be evaluated to characterize the uncertainty about the hypotheses *A*. The belief function, *Bel*, measures the minimum uncertainty value about *A*, whereas, the plausibility function, *Pls*, reflects the maximum uncertainty value. Belief and plausibility functions are defined from 2^{Θ} to [0, 1]

$$Bel(A) = \sum_{A_i \subseteq A} m(A_i), \quad Pls(A) = \sum_{A_i \cap A \neq \phi} m(A_i) \qquad (\mathbf{3})$$

2.2 The Dezert-Smarandache Theory (DSmT)

While the DST considers Θ as a set of exclusive elements, the DSmT relaxes this condition and allows for overlapping and intersecting hypotheses. This allows for quantifying the conflict that might arise between the different sources throughout the assignment of non-null confidence values to the intersection of distinct hypotheses.

Let $\Theta = \{\theta_1, \theta_2, \ldots, \theta_N\}$be a set of *N* elements which can potentially overlap. The hyper-power set D^{Θ} is defined as the set of all composite hypotheses obtained from Θ with $\cup$ and $\cap$ operators such that

1. $\phi, \theta_1, \theta_2, \ldots, \theta_N \in D^{\Theta}$.
2. If A, B $\in$ D^{Θ}, then $(A \cup B) \in D^{\Theta}$ and $(A \cap B) \in D^{\Theta}$.
3. No other elements belong to D^{Θ}, except those defined in 1) and 2).

As in the DST, the DSmT defines a map m(.): $D^{\Theta} \rightarrow [0,1]$. This map defines the confidence level that each sensor associates with the element of D^{Θ}. This map supports paradoxical information, and $\sum_{A \in D^{\Theta}} m(A) = 1$.

(c) Department of Computer Science, Bishop's University
http://cs.ubishops.ca/research.html

The belief and plausibility functions are defined in the same way as for the DST. The DSmT rule of combination of conflicting and on uncertain sources is given by

$$\sum_{A \in D^{\Theta}} m(A) = 1 \tag{4}$$

and

$$m(A) = \sum_{\substack{A_1, A_2, \ldots, A_d \in D^{\Theta} \\ A_1 \cap A_2 \cap \ldots \cap A_d = A}} \prod_{i=1}^{D} m_i(A_i) \,. \tag{5}$$

3 Sequential Monte Carlo

Sequential Monte Carlo techniques, also known as particles filtering, were introduced to track multiple objects in cluttered scenes [6]. In the following, let's consider $X_t = (x_1, x_2, \ldots, x_t)$ as the state vector (location, size, etc.) describing the target and $Z_t = (z_1, z_2, \ldots, z_t)$ as the vector of measurements (color, texture, etc.) up to time t. The tracking is based on the estimation of the posterior state distribution $p(x_t|Z_t)$ at each time step. The estimation is performed using a two step Bayesian recursion. The first step is prediction,

$$p(x_t \mid Z_{t-1}) \propto \int p(x_t \mid x_{t-1}) \, p(x_{t-1} \mid Z_{t-1}) dx_{t-1} \tag{6}$$

and the second step is filtering

$$p(x_t \mid Z_t) \propto p(z_t \mid x_t) \, p(x_t \mid Z_{t-1}) \tag{7}$$

This recursion requires the specification of the state evolution $p(x_t|x_{t-1})$ and a measurement model linking the state and the current measurement $p(z_t|x_t)$. The basic idea behind the particles filtering is very simple. Starting with a weighted set of samples

$$S_{t-1} = \{ s_{t-1}^{(n)}, \pi_{t-1}^{(n)} \mid \sum_{n=1}^{N} \pi_{t-1}^{(n)} = 1 \} \tag{8}$$

which describe target candidates and distributed according to $p(x_{t-1}|z_{t-1})$, new samples are obtained by propagating each sample according to the target's state model, $p(x_t|x_{t-1})$. In the filtering step, each sample is weighted given the observation and N samples are drawn with replacement according to $\pi_t = p(z_t|x_t)$. The value given by $E[S_t] = \sum_{n=1}^{N} \pi_t^{(n)} s_t^{(n)}$ will represent the best estimation of the target. Particles filtering has proven to be very successful for non-linear and non-Gaussian estimation problems. The tracking iteration using particles filtering can be summarized as follows.

(c) Department of Computer Science, Bishop's University
http://cs.ubishops.ca/research.html

<u>Step 1</u>: Select N samples_

$$S_{t-1} = \{s_{t-1}^{(n)}, \pi_{t-1}^{(n)} \mid \sum_{n=1}^{N} \pi_{t-1}^{(n)} = 1\} \quad \textbf{(9)}$$

<u>Step 2</u>: Propagate each sample according to

$$s_t^{(n)} = A \cdot s_{t-1}^{(n)} + w_{t-1}^{(n)} \quad \textbf{(10)}$$

where A is a square matrix defining the determistic component of the target's motion model and $w_{t-1}^{(n)}$ is the random component of the target's motion model.

<u>Step 3</u>: Observe the samples and calculate $\pi_t^{(n)}$.

<u>Step 4</u>: Estimate the mean state of S_t

$$E[S_t] = \sum_{n=1}^{N} \pi_t^{(n)} s_t^{(n)} \quad \textbf{(11)}$$

4 DSmT-Based Tracking

In this section, we describe a general framework for multiple target tracking using multiple cue. This approach uses the DSmT combinational rule to refer the information provided by the cues into a single representation. This latter takes into account the conflicts between the cues that might arise due to occlusion. Let's assume that the number of targets, τ, and the number of cues, c, are known. Up to time t-1, each target is associated with a track $\{\theta_j\}_{j=1}^{\tau}$. At time t, an image frame is extracted from the video sequence and a number of measurements are obtained for each target candidate. Thus, the objective is to combine these measurements in order to determine the best track for each candidate. It is important to notice that a target candidate, in this paper, refers to a particle sample. The hyper-power set D^{Θ} defines the set of the hypotheses for which the different cues can provide confidence values. These hypotheses can correspond to: 1) individual tracks θ_j, 2) union of tracks $\theta_r \cup \ldots \cup \theta_s$, which symbolizes ignorance, 3) intersection of tracks $\theta_r \cap \ldots \cap \theta_s$, which symbolizes conflict or 4) any tracks combination obtained by $\cup$ and $\cap$ operators. The confidence level is expressed in terms of mass function $\{m_{t,l}^{(n)}(.)\}_{l=1}^{c}$ that is committed to each hypothesis and which satisfies the condition in (4). Given this framework, $m_{t,l}^{(n)}(A)$ expresses the confidence with which cue l associates particle n to hypothesis A at time t. According to DSmT combinational rule in (6), a single map function $m_t^{(n)}(.)$ can be derived as follows

$$m_t^{(n)}(A) = m_{t,1}^{(n)}(.) \oplus m_{t,2}^{(n)}(.) \oplus \cdots \oplus m_{t,c}^{(n)}(.) \quad \textbf{(12)}$$

(c) Department of Computer Science, Bishop's University
http://cs.ubishops.ca/research.html

where $m_t^{(n)}(A)$ is the overall confidence level with which all cues associate particle n to hypothesis A at time t. Since the target candidates must be associated to individual tracks, the information contained in compound hypotheses is transferred into single hypotheses (i.e. single tracks) through the notions of the belief or plausibility functions

$$Bel_t^{(n)}(\theta_j) = \sum_{\substack{\theta_i \subseteq A \\ A \in D^\Theta}} m_t^{(n)}(A), Pls_t^{(n)}(\theta_j) = \sum_{\substack{\theta_i \cap A \neq \phi \\ A \in D^\Theta}} m_t^{(n)}(A) \quad \textbf{(13)}$$

Where $Bel_t^{(n)}(\theta_j)$ (resp. $Pls_t^{(n)}(\theta_j)$) quantifies the confidence with which particle n is associated to θ_j at time t using the notion of belief (resp. plausibility). The confidence levels in (13) are not used to determine whether a given candidate is the best estimate or not of the target, they are rather used to quantify the weight of the candidate (or particle) as a sample of the state posterior distribution $p(x_t|Z_t)$. The DSmT-based particles filtering algorithm implemented in this paper is given below.

Step 1: Initialization

- generate N samples $S_{t,j} = \{s_{t,j}^{(n)}, \pi_{t,j}^{(n)}\}_{n=1}^N$ for each target $j = 1,\ldots,\tau$ independently, with $\pi_{t,j}^{(n)} = \frac{1}{N}$. Set t =1.

Step 2: Propagation

- $s_{t*,j}^{(n)} = A \cdot s_{t-1,j}^{(n)} + w_{t-1,j}^{(n)}$

Step 3: Observation (for each particle)

- Compute $\{m_{t*,l}^{(n)}(A)\}_{l=1}^c$ for $A \in D^\Theta$
- Compute $m_{t*}^{(n)}(A)$ for $A \in D^\Theta$ according to (5)
- Calculate the particle weight $\pi_{t*,j}^{(n)} = Bel_{t*}^{(n)}(\theta_j)$ (or $\pi_{t*,j}^{(n)} = Pls_{t*}^{(n)}(\theta_j)$)
- Normalize the weight: $\tilde{\pi}_{t*,j}^{(n)} = \frac{\tilde{\pi}_{t*,j}^{(n)}}{\sum_{n=1}^N \tilde{\pi}_{t*,j}^{(n)}}$

Step 4: Estimation

- Target $j = 1,\ldots,\tau$ is given by $E[S_{t*,j}] = \sum_{n=1}^N \tilde{\pi}_{t*,j}^{(n)} s_{t*,j}^{(n)}$.

Step 5: Resampling (for each target)

- Generate $S_{t,j} = \{s_{t,j}^{(n)}, \pi_{t,j}^{(n)}\}_{n=1}^N$ by resampling N times from $S_{t*,j}$ where $p(s_{t,j}^{(n)} = s_{t*,j}^{(n)}) = \tilde{\pi}_{t*,j}^{(n)}$.

Step 6: Incrementing

- $t = t + 1$, go to step 2.

5 Tracking Two Targets Using Location and Color

For two targets, we can define Θ as follows

(c) Department of Computer Science, Bishop's University
http://cs.ubishops.ca/research.html

$$\Theta = \{\theta_1, \theta_2, \overline{\theta_1 \cup \theta_2}\} \quad \textbf{(14)}$$

In (14), θ_1 refers the first target, θ_2 refers the second target and $\overline{\theta_1 \cup \theta_2}$ refers the rest of the scene. Actually, hypothesis $\overline{\theta_1 \cup \theta_2}$ **c**an refer the background information. However, since this latter can change during the tracking, we will refer to $\overline{\theta_1 \cup \theta_2}$ as the false alarm hypothesis. Beside, $\theta_1 \cap \theta_2 \neq \phi$ due to the possible occlusion, and $\theta_j \cap \overline{\theta_1 \cup \theta_2} = \phi$ for j=1,2.

5.1 The Location Cue

The targets locations at time t-1 are known and given by $(x_{t-1,1}, y_{t-1,1})$ and $(x_{t-1,2}, y_{t-1,2})$. At time t, the probability that a particle $s_{t,j}^{(n)}$ located at $(x_{t,j}^{(n)}, y_{t,j}^{(n)})$ belongs to target j=1,2 according to the location information is defined from a Gaussian pdf as follows

$$p_{t,j}^{(n)} = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_{t,j}^{(n)} - x_{t-1,j})^2 + (y_{t,j}^{(n)} - y_{t-1,j})^2}{2\sigma^2}} \quad \textbf{(15)}$$

where σ is a bandwidth parameter. Similarly, the probability that a given particle does not belong to θ_1 and θ_2 is inversely proportional to the distance between the particle and both targets. Since Θ is exhaustive, a particle that does not belong to θ_1 and θ_2 do belong to $\overline{\theta_1 \cup \theta_2}$. This leads us to the definition of a new pdf, $p_{t,FA}^{(n)}$, which measures the membership of a particle n = 1,…,N to the false alarm hypothesis

$$p_{t,FA}^{(n)} = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(d_{\max} - d_{1-2}^{(n)})^2}{2\sigma^2}} \quad \textbf{(16)}$$

where σ is the bandwidth parameter, d_{max} is the radius of a circle centered on the mid-point of targets 1 and 2, and which contains all the particles used for tracking at the time t-1, $d_{1-2}^{(n)}$ is the distance separating particle n and the mid-point.

$$d_{1-2}^{n} = \sqrt{(x_{t,1}^{(n)} - \frac{x_{t-1,1} - x_{t-1,2}}{2})^2 + (y_{t,1}^{(n)} - \frac{y_{t-1,1} - y_{t-1,2}}{2})^2} \quad \textbf{(17)}$$

The mass function of particle n according to its location is given by

$$m_{t,1}^{(n)}(\theta_j) = \frac{p_{t,j}^{(n)}}{p_{t,1}^{(n)} + p_{t,2}^{(n)} + p_{t,FA}^{(n)}}, \quad j\text{=1, 2} \quad \textbf{(18)}$$

(c) Department of Computer Science, Bishop's University
http://cs.ubishops.ca/research.html

$$m_{t,1}^{(n)}(\overline{\theta_1 \cup \theta_2}) = \frac{p_{t,FA}^{(n)}}{p_{t,1}^{(n)} + p_{t,2}^{(n)} + p_{t,FA}^{(n)}} \quad \textbf{(19)}$$

5.2 The Color Cue

Let's assume that both target models are known and given by normalized color histograms $\{q_j(u)\}_{u=1}^{m}$, where u is a discrete color index and m is the number of histogram bins. At time t, the normalized color histogram of particle $s_{t,j}^{(n)}$ is given by $\{h_{t,j}^{(n)}(u)\}_{u=1}^{m}$. The probability that particle $s_{t,j}^{(n)}$ belongs to target j=1,2 according to the color histogram is derived from the following Gaussian pdf.

$$p_{t,j}^{(n)} = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(d_{t,j}^{(n)})^2}{2\sigma^2}} \text{, } j=1,2 \quad \textbf{(20)}$$

where σ is a color bandwidth parameter, $d_{t,j}^{(n)}$ is the Bhattacharyya distance between $h_{t,j}^{(n)}(u)$ and $q_j(u)$ at time t

$$d_{t,j}^{(n)} = \sqrt{1 - \sum_{u=1}^{m} h_{t,j}^{(n)}(u) q_j(u)} \quad (21)$$

Let's define $\{q_{FA}(u)\}_{u=1}^{m}$ as the histogram of the scene from which we subtract the histogram of targets 1 and 2.

$$q_{FA}(u) = \max\{q_{scene}(u) - q_1(u) - q_2(u), 0\} \quad \textbf{(22)}$$

The probability that $s_{t,j}^{(n)}$ belongs to the false alarm hypothesis will be given by

$$p_{t,FA}^{(n)} = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(d_{t,FA}^{(n)})^2}{2\sigma^2}} \quad \textbf{(23)}$$

where

$$d_{t,FA}^{(n)} = \sqrt{1 - \sum_{u=1}^{m} h_{t,j}^{(n)}(u) q_{FA}(u)} \text{ .} \quad \textbf{(24)}$$

The mass functions of particle n according to color can be evaluated as follows

$$m_{t,2}^{(n)}(\theta_j) = \frac{p_{t,j}^{(n)}}{p_{t,1}^{(n)} + p_{t,2}^{(n)} + p_{t,FA}^{(n)}} \text{, } j=1,2 \quad \textbf{(25)}$$

(c) Department of Computer Science, Bishop's University
http://cs.ubishops.ca/research.html

$$m_{t,2}^{(n)}(\overline{\theta_1 \cup \theta_2}) = \frac{p_{t,FA}^{(n)}}{p_{t,1}^{(n)} + p_{t,2}^{(n)} + p_{t,FA}^{(n)}} \quad \textbf{(26)}$$

5.3 The Cue Combination

The combination rule leads to the mass functions $m_t^{(n)}(.)$ that are defined in table-1

Table 1. Combination rule for color and location

		Color Cue		
		$m_{t,2}^{(n)}(\theta_1)$	$m_{t,2}^{(n)}(\theta_2)$	$m_{t,2}^{(n)}(\overline{\theta_1 \cup \theta_2})$
Location Cue	$m_{t,1}^{(n)}(\theta_1)$	$m_t^{(n)}(\theta_1)$	$m_t^{(n)}(\theta_1 \cap \theta_2)$	ϕ
	$m_{t,1}^{(n)}(\theta_2)$	$m_t^{(n)}(\theta_1 \cap \theta_2)$	$m_t^{(n)}(\theta_2)$	ϕ
	$m_{t,1}^{(n)}(\overline{\theta_1 \cup \theta_2})$	ϕ	ϕ	$m_t^{(n)}(\overline{\theta_1 \cup \theta_2})$

where

$$m_t^{(n)}(\theta_1) = m_{t,1}^{(n)}(\theta_1) \cdot m_{t,2}^{(n)}(\theta_1) \quad \textbf{(27)}$$

$$m_t^{(n)}(\theta_2) = m_{t,1}^{(n)}(\theta_2) \cdot m_{t,2}^{(n)}(\theta_2) \quad \textbf{(28)}$$

$$m_t^{(n)}(\theta_1 \cap \theta_2) = m_{t,1}^{(n)}(\theta_1) \cdot m_{t,2}^{(n)}(\theta_2) + m_{t,1}^{(n)}(\theta_2) \cdot m_{t,2}^{(n)}(\theta_1) \quad \textbf{(29)}$$

$$m_t^{(n)}(\overline{\theta_1 \cup \theta_2}) = m_{t,1}^{(n)}(\overline{\theta_1 \cup \theta_2}) \cdot m_{t,2}^{(n)}(\overline{\theta_1 \cup \theta_2}) \quad \textbf{(30)}$$

$$m_t^{(n)}(\phi) = m_{t,1}^{(n)}(\overline{\theta_1 \cup \theta_2})(m_{t,2}^{(n)}(\theta_1) + m_{t,2}^{(n)}(\theta_2)) + m_{t,2}^{(n)}(\overline{\theta_1 \cup \theta_2})(m_{t,1}^{(n)}(\theta_1) + m_{t,1}^{(n)}(\theta_2)) \quad \textbf{(31)}$$

Equation (27) is the confidence level with which both cues associate $s_{t,j}^{(n)}$ to target 1. Equation (28) is the confidence level with which both cues associate $s_{t,j}^{(n)}$ to target 2. Equation (29) is the conflict value between the cues for the membership of $s_{t,j}^{(n)}$ to target 1 or target 2. Equation (30) expresses the confidence value with which both cues agree that the particle corresponds to a false alarm. Equation (31) quantifies the conflict between the targets and the false alarm hypothesis.

The weight of particle $s_{t,j}^{(n)}$ within the posterior $p(X_t|Z_t)$ distribution, for target j, is calculated using belief (or the plausibility) function, so that

(c) Department of Computer Science, Bishop's University
http://cs.ubishops.ca/research.html

$$\pi_{t,j}^{(n)} = Bel_t^{(n)}(\theta_j) = m_t^{(n)}(\theta_j) + m_t^{(n)}(\theta_1 \cap \theta_2), \quad j=1,2 \qquad \textbf{(32)}$$

The generalization of the tracking scheme described in this section to τ targets can be carried out by defining a frame of discernment $\Theta = \{\theta_1, \cdots, \theta_\tau, \overline{\theta_1 \cup \cdots \cup \theta_\tau}\}$, where θ_j are individual targets and $\overline{\theta_1 \cup \cdots \cup \theta_\tau}$ is the false alarm hypothesis.

6 Experimental Results

The tracking approach is tested on a cluttered scene with intersecting targets. Two targets are selected from frame 1 as shown in figure 1. Since the purpose of this paper is tracking and not the target detection, both targets are selected manually. Each target is tracked using 20 particles only. An increased number of particles will result in a smoother tracking, while increasing the processing time. For the sake of clarity, the person on the right in the first image is denoted target 1 and the person on the left is denoted target 2.

Fig. 1. First row: Tracking during the pre-occlusion phase. **Second row**: Tracking during the occlusion phase. **Third row:** Tracking during the post-occlusion phase.

The tracking sequence is divided into three phases. Phase 1 is the pre-occlusion sequence, phase 2 corresponds to the occlusion sequence, and phase 3 is the post-occlusion sequence. Tracking in phase 2 is challenging due to the closeness of the targets, which perturbs the measured cues and might lead to a false identification.

(c) Department of Computer Science, Bishop's University
http://cs.ubishops.ca/research.html

The location cue loses gradually its ability to separate targets 1 and 2 as they converge to the intersection point. However, the location cue remains a valid measurement because it is independent from the relative location of targets with respect to the camera (occluding or occluded). The color cue is extremely sensitive to the occlusion. During phase 2, target 1 is partially or totally occluded by target 2. As a result, the color measurement for particles associated with target 1 is corrupted by the presence of target 2. When the occlusion is total, target 1 disappears from the scene and the color measurement becomes invalid. The occlusion affects also the behavior of particles associated with target 2 since the presence of target 1 in its neighborhood will be interpreted by the algorithm as a rapid change in the background information. The tracking performances in phase 3 depend on the outcome of tracking during phase 2. A successful tracking would result on a correct identification; whereas, a failure would result on a bad identification of the occluded target (target 1). As shown in figure 1, the approach proposed in this paper accurately identifies the targets during the three phases of the tracking. This is due to the effective handling of the conflicting information provided by the location and color cues during the second phase of tracking by the DSmT model. Figure 2 shows the variation of the average value of the confidence levels for all particles during the tracking.

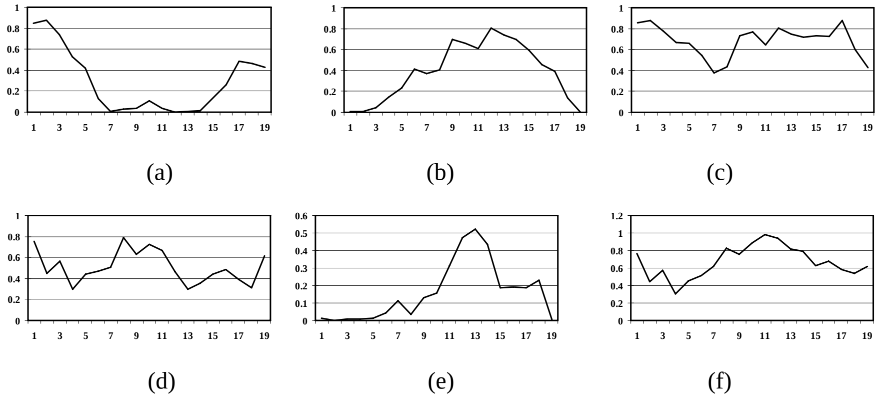

Fig. 2. The variation of: a) $m_{avg}(\theta_1)$, b) $m_{avg}(\theta_1 \cap \theta_2)$, c) $Bel_{avg}(\theta_1)$ for the occluded target. The variation of : d) $m_{avg}(\theta_2)$, e) $m_{avg}(\theta_1 \cap \theta_2)$, f) $Bel_{avg}(\theta_2)$ for the occluding target

$$m_{avg}(\theta_j) = \frac{1}{N}\sum_{n=1}^{N} m_t^{(n)}(\theta_j), \quad m_{avg}(\theta_1 \cap \theta_2) = \frac{1}{N}\sum_{n=1}^{N} m_t^{(n)}(\theta_1 \cap \theta_2), \text{ and}$$

$$Bel_{avg}(\theta_j) = \frac{1}{N}\sum_{n=1}^{N} Bel_t^{(n)}(\theta_j).$$

The confidence level for the occluded target, $m_{avg}(\theta_1)$, is high during phases 1 and 3, but it decreases in phase 2 (see figure 2-a). Indeed, in phases 1 and 3 the color and location cues both agree on the identity of the target. However, in phase 2 the target is occluded and this reduces the confidence value provided by the color cue.

During the same phase, the location confidence remains high, which explains the increase in the conflict, $m_{avg}(\theta_1 \cap \theta_2)$, as shown in figure 2-b. The belief function

(c) Department of Computer Science, Bishop's University
http://cs.ubishops.ca/research.html

$Bel_{avg}(\theta_1)$ is given in figure 2-c. This curve shows the high confidence with which the target is located despite the occlusion. This is mainly due to the introduction of the conflict information through the DSmT model. Figures 2-d, 2-e and 2-f show that the effect of the occlusion on the occluding target is small in comparison with its effect on the occluded target. The existence of such an effect can be justified by the presence of target 1 in the immediate neighborhood of target 2, which rapidly modifies the color measurement for some particles.

7 Conclusion

In this paper, we addressed the problem of tracking multiple target in a cluttered scene using multiple cue. Within this framework, we developed a model that combines the classical particles filtering approach and the novel DSmT theory. A set of particles is used to track each target. The DSmT model assigns confidence level values for the membership of each particle. This membership takes into consideration the conflict between the cues during the occlusion phase, allowing thus a better tracking. The experimental results demonstrated the effectiveness of the model in case of multiple target tracking using location and color and its interest in cluttered scenes. We also showed how our approach can easily be generalized to deal with additional cues and targets.

References

1. C. Chen, X. Lin, and Y. Shi.: Moving object tracking under varying illumination conditions. Pat. Rec. Let., Vol.27 (2006) 1632-1643
2. E. Ozyilidiz, N. Krahnstover, R. Sharma.: Adaptive texture and color segmentation for tracking moving objects. Pat. Rec., Vol. 35, No. 10. (2002) 2013-2029
3. B. McCane, B. Galvin, and K. Novins.: Algorithmic fusion for more robust feature tracking. Int. J. of Com. Vis., Vol. 49, No. 1, (2002) 79-89
4. G. Shafer.: A Mathematical Theory of Evidence. Princeton University Press, (1976)
5. F. Smarandache and J. Dezert.: Applications and Advances of DSmT for Information Fusion. Am. Res. Press, Rehoboth (2004)
6. M.S. Arulampalam, S. Maskell, N. Gordon and T. Clapp.: A Tutorial on Particles filterings for Online Nonlinear/Non-Gaussian Bayesian Tracking. IEEE Tran. On Sig. Proc., Vol. 50, No. 2. (2002) 174-1

(c) Department of Computer Science, Bishop's University
http://cs.ubishops.ca/research.html