Research Journal of Applied Sciences, Engineering and Technology 6(12): 2131-2135, 2013
ISSN: 2040-7459; e-ISSN: 2040-7467

Submitted: December 3, 2012 Accepted: January 17, 2013 Published: July 30, 2013

Visual Multi-Object Tracking in the Presence of Cluttered Scenes

[1]Xu-Sheng Gan and [2]Yi Wang
[1]XiJing College, Xi'an 710123, Shaanxi China
[2]Air Force Engineering University, Xi'an 710077, Shaanxi China

Abstract: The aim of this study was to investigate the visual multi-object tracking in the presence of cluttered scenes. A improved algorithm of fusing multi-source information including location and color evidences were introduced based on Dezert-Smarandache Theory (DSmT) and Particle Filters (PF). Results showed that the conflict strategy and DSmT combination model were available and the proposed approach exhibited a significantly better performance for dealing with high conflict between evidences than a conventional PF. The suggested approach can easily be generalized to deal with larger number of visual multi-object and additional cues in the presence of cluttered scenes.

Keywords: Combination method, DSmT, multi-object, PF

INTRODUCTION

Research on multi-object tracking in the part decades has yielded an arsenal of powerful algorithms, which play important roles in many applications, such as visual surveillance, robotics and human-computer interaction, etc (Zhao and Nevatia, 2004; Han *et al*., 2011). Hence, Visual multi-object tracking is one of the research hotspots in the field of computer vision. Although many effective visual multi-object tracking methods have been proposed, there are still a lot of difficulties in designing a robust tracking algorithm due to the challenging complex scenarios such as significant illumination changes in natural environment, pose variations of the object and non-linear deformations of shapes and noise and dense clutters in complex background (Kazuhiro, 2009).

In the last fifteen years we have witnessed a rapid development of the theory of Particle Filters (PF) and the corresponding algorithms of PF are widely applied in tracking fields (Lefevre, 2002; Abdallah *et al*., 2008; Crisan and Obanubi, 2012). Although lots of algorithms have been introduced to track moving multi-object in different cases, these approaches have mainly improved local performances by optimizing PF algorithms and there still exist many key issues which need to be discussed further. Recently, Dezert-Smarandache Theory (DSmT) by Dezert and Smarandache (2004) has to be viewed as a general flexible bottom-up approach for managing uncertainty and conflicts for a wide class of static or dynamic fusion problems where the information to combine is modeled as a finite set of belief functions provided by different independent sources of evidence (Tehamova *et al*., 2005). Hence, the DSmT of plausible and paradoxical reasoning has become very important method to deal with highly conflicting, uncertain and imprecise sources of evidence in intelligent system by information fusion. The corresponding researches showed the conflicting focal elements were increased based on DSmT and the computational effort was also increased in the course of reasoning. Based on this matter of fact, improved methods were presented to reduce the computational efforts (Yang *et al*., 2010). At present, some researches focused on visual multi- object identification based on DSmT, the related studies of visual multi- object tracking based on DSmT are very less.

How to develop a robust and real-time visual multi-object tracking approach in the presence of cluttered scenes is very necessary. The aim of this study is to present a novel approach of visual multi-object tracking that will handle the multi-object of crosses and occlusions in order to attain an excellent information fusion.

COMBINATION METHOD

PF has been proven to be very successful for nonlinear and non-Gaussian estimation problems and the basic tracking step includes selection of samples, propagation of samples, observation of samples and calculation and estimation of the mean state and the detailed content can be found in literatures (Djuric *et al*., 2003; Lao *et al*., 2009). the following addressed the tracking approach based on DSmT in detail.

If the number of objects is τ, the number of cues is c and the τ and c are known. Up to time t-1, each object is associated with a track $\left\{\{\theta_j\}_{j=1}^{\tau}\right\}$. At time t, an image frame is extracted from the video sequence and a number of measurements are obtained for each object

Corresponding Author: Yi Wang, Air Force Engineering University, Xi'an 710077, Shaanxi China

candidate. Thus, the object given is to combine these measurements in order to determine the best track for each candidate. It is important to notice that a object candidate, in this study, 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:

- Individual tracks θ_j
- Union of tracks $\theta_r \cup \ldots \cup \theta_s$, which symbolizes ignorance
- Intersection of tracks $\theta_r \cap \ldots \cap \theta_s$, which symbolizes conflict or
- Any tracks combination obtain by $\cup$ and $\cap$ operators

The confidence level is expressed in terms of mass function $\{m^{(n)}_{t,l}(.)\}^{c}_{l=1}$ that is committed to each hypothesis and which satisfies the condition in (4). Given this framework, $m^{(n)}_{t,l}$ (A) expresses the confidence with which cue l associates particle n to hypothesis A at time t. A single map function $m^{(n)}_{t}(.)$can be derived as follows based to DSmT combinational rule:

$$m_t^{(n)}(A) = m_{t,1}^{(n)}(\cdot) \oplus m_{t,2}^{(n)}(\cdot) \oplus \cdots \oplus m_{t,c}^{(n)}(\cdot) \tag{1}$$

where, $m^{(n)}_{t}(A)$ denotes the overall confidence level with which all cues associate particle n to hypothesis A at time t.

Since the object 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, is given by Brun *et al*. (2002) and Marcelo (2005):

$$Bel_t^{(n)}(\theta_j) = \sum_{\substack{\theta_i \subseteq A \\ A \in D^{\Theta}}} m_t^{(n)}(A) \tag{2}$$

$$Pls_t^{(n)}(\theta_j) = \sum_{\substack{\theta_i \cap A \neq \Phi \\ A \in D^{\Theta}}} m_t^{(n)}(A) \tag{3}$$

where, $Bel_t^{(n)}(\theta_j)$ (resp. $Pls^{(n)}_{t}(\theta_j)$ 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 are not used to determine whether a given a candidate is the best estimate or not of the object, they are rather used to quantify the weight of the candidate as a sample of the state posterior distribution $p(X_t|Z_t)$. As a result, the PF algorithm based on DSmT can be implemented.

Because conflicting focal elements preserved can increase assignment of the focal element in the framework of DSmT, the convergence is very slow for assign function of main focal element in most cases and the difficulty of tracking is increased greatly. In this study, a modified conflict strategy was presented. Namely, local conflict not was assigned globally but was assigned locally by refining global conflict into r local conflicts.

Based on the above analysis, the following established dynamic combination model of multi-object tracking. In order to describe conveniently, the cues of color and location were used to track two objects. For two objects, Θ was defined as follows:

$$\Theta = \{\theta_1, \theta_2, \overline{\theta_1 \cup \theta_2}\} \tag{4}$$

In Eq. (4), θ_1 refers the first object, θ_2 refers to the second object and $\overline{\theta_1 \cup \theta_2}$ refers to the rest of the scene. Actually, hypothesis $\overline{\theta_1 \cup \theta_2}$ **c**an refer to the background information. Since this latter can change during the tracking, we will refer to $\overline{\theta_1 \cup \theta_2}$ as the false alarm hypothesis. Beside, $\overline{\theta_1 \cap \theta_2} = \emptyset$ due to the possible occlusion and $\theta_j \cap \overline{\theta_1 \cap \theta_2} = \emptyset$ for j=1,2.

Based on the modified combination strategy introduced, the combination rule leads to the mass function $m_t^{(n)}(.)$ and the corresponding combination rules of color and location are defined as follows:

$$m_t^{(n)}(\theta_1) = m_{t,1}^{(n)}(\theta_1) \cdot m_{t,2}^{(n)}(\theta_1) \tag{5}$$

$$m_t^{(n)}(\theta_2) = m_{t,1}^{(n)}(\theta_2) \cdot m_{t,2}^{(n)}(\theta_2) \tag{6}$$

$$m_t^{(n)}(\theta_1 \cap \theta_2) = m_{t,1}^{(n)}(\theta_1) \cdot m_{t,2}^{(n)}(\theta_2) + m_{t,1}^{(n)}(\theta_2) \cdot m_{t,2}^{(n)}(\theta_1) \tag{7}$$

$$m_t^{(n)}(\overline{\theta_1 \cup \theta_2}) = m_{t,1}^{(n)}(\overline{\theta_1 \cup \theta_2}) \cdot m_{t,2}^{(n)}(\overline{\theta_1 \cup \theta_2}) \tag{8}$$

$$\begin{aligned} m_t^{(n)}(\phi) &= m_{t,1}^{(n)}(\overline{\theta_1 \cup \theta_2})(m_{t,2}^{(n)}(\theta_1) + m_{t,2}^{(n)}(\theta_2)) \\ &+ m_{t,2}^{(n)}(\overline{\theta_1 \cup \theta_2})(m_{t,1}^{(n)}(\theta_1) + m_{t,1}^{(n)}(\theta_2)) \end{aligned} \tag{9}$$

Eq. (5) is the confidence level with which both cues associate $S_{t,j}^{(n)}$ to object 1. Eq. (6) is the confidence level with which both cues associate $S_{t,j}^{(n)}$ to object 2. Eq. (7) is the conflict value between the cues for the membership of $S_{t,j}^{(n)}$ to object 1 or object 2. Eq. (8) expresses the confidence value with which both cues agree that the particle corresponds to a false alarm. Eq. (9) quantifies the conflict between the objects and the false alarm hypothesis.

MULTI-OBJECT TRACKING

Tracking test: A tracking under occluding conditions was tested based on crossing objects and closely spaced objects. Two objects were selected in

Table 1: Two crossing objects values of data association

	Values of data association			
Object tracking	Tracking 1	Tracking 2	Missing	Wrong
Object 1	0.774	0.120	0.014	0.007
Object 2	0.135	0.768	0.014	0.003

Table 2: Two closely spaced objects values of data association

	Values of data association			
Object tracking	Tracking 1	Tracking 2	Missing	Wrong
Object 1	0.665	0.235	0.002	0.0004
Object 2	0.203	0.450	0.001	0.0000

order to analyze conveniently and each object was tracked using 20 particles only. The tracking sequence was divided into three phases. Phase 1 was the pre-occlusion sequence, phase 2 corresponded to the occlusion sequence and phase 3 was the post-occlusion sequence. Tracking in phase 2 was challenged due to the closeness of the object, which perturbs the measured cues and might led to a false identification. The simulation scenario consisted of two objects and a stationary sensor was at the origin with T = 5 sec, measurement standard deviations 0.4° and 70 m for azimuth and range, respectively.

Firstly, for two crossing moving objects, the tracking algorithm based on DSmT was applied to produce the attribute probability term in generalized assignment matrix and the corresponding computing results of tracks' purity in case of generalized data association was gotten as shown in Table 1. Where missing was used for the case when in the track's gate there was no observation and wrong was used for the case when the track was associated with the false alarm. It was very obvious that the tracks' purity was increased from the result in Table 1. Secondly, two closely spaced objects will be analyzed. For two closely spaced objects and one could easily see that the two closely spaced moving in parallel objects lost the proper directions and the tracks switch. At last, the tracking algorithm based on DSmT was applied to discuss the data association problem and the corresponding results based on DSmT was given in Table 2. The results showed the proper data associations in condition of two closely spaced objects.

Tracking example: Further, a video scene including two pedestrians was a campus region and the tracking experiment of only two objects in natural environment was carried out to validate the proposed tracking approach. The video was captured from internet platforms. In the tracking experiment, image pre-processing was employed and the initial position of objects was manual designated. Initialization of tracking was executed at the beginning of every image subsequence, which included calibration the location and space area of objects tracked, estimation the moving direction and speed of every object and calculation the scaling according to the trend of relative motion between objects tracked and imaging lens. Let the two pedestrians in video scene keep uniform motion along the moving direction of cross and occlusion and the illumination changes be also omitted.

In the process of tracking, the color distribution of objects had obviously difference by comparing with the surrounding environment and the surface feature relative to the distributed location of structural distortion is very much small. At the same, there existed high conflict problems including scale variations, cross and occlusion of objects. From 60 to 81 frames, the two objects underwent the following processes including cross, partly and fully occlusion. The number of particle was variable with conflict levels between evidences and the maximum number of particles was 40 when the two objects were basically covered. Namely, 40 particles were only used to handle the high conflict between evidences. Finally, the whole video-based tracking was accomplished and the video demonstrating tracking results were available by the proposed approach. Figure 1 shows the main frames with tracking particles during tracking experiment and

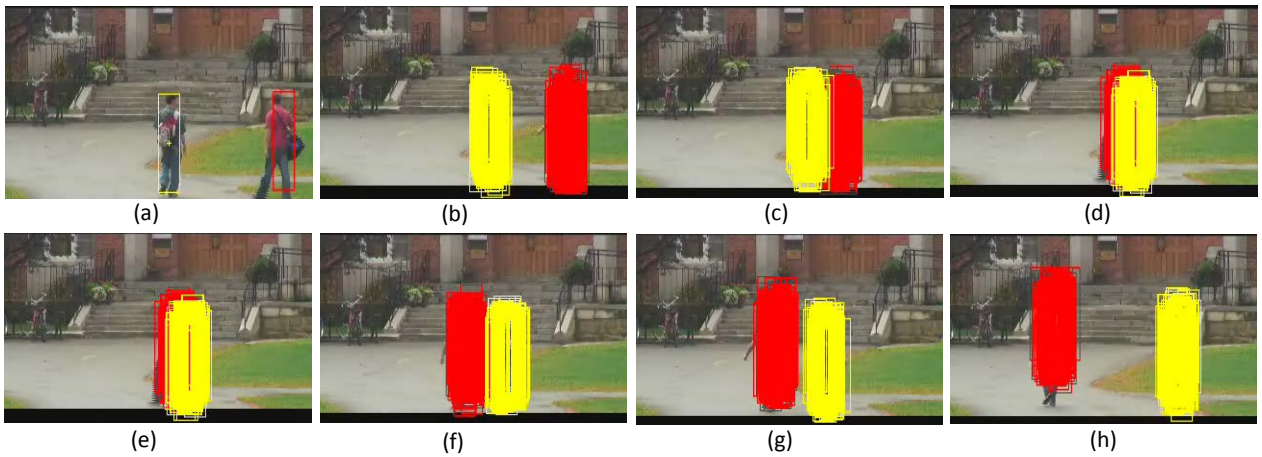

Fig. 1: Main frames with tracking particles during tracking

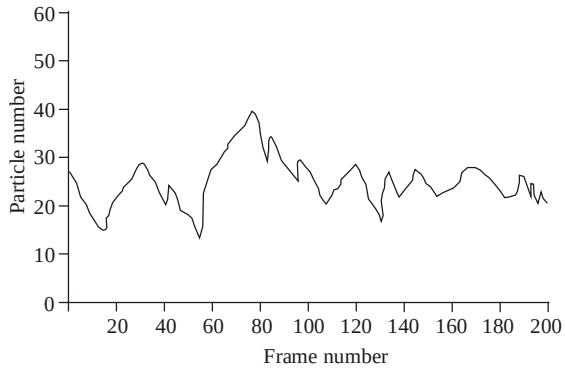

Fig.2 Variation of particle number in different tracking stages

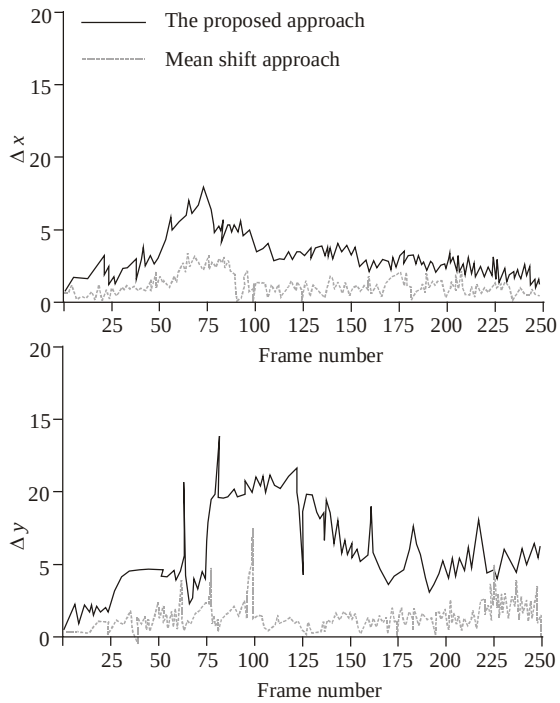

Fig. 3: Deviation of object center of the two approaches

Fig. 2 shows the variation of particle number in different multi- object tracking stages.

In order to validate the availability and stability for handling high conflict between evidences by the proposed approach, a mean shift approach was also applied to track the above image sequence. At last, the tracking process and result of two kinds of approach were obtained. Figure 3 shows the deviation of object center (Δx, Δy) during tracking by comparing with the two approaches.

It was seen from Fig. 3 that the variation of deviation of object center was much smaller than that of the mean shift approach during almost the whole tracking. However, the mean shift approach's accuracy deteriorates rather quickly when the two objects had high cross and occlusion. Thereby, the results showed that the suggested approach could track visual multi-object effectively and the approach had better adaptation to object and background variation. According to the whole tracking course, the proposed approach accurately identifies different visual object during different tracking stages. This is due to the effective handling of the conflicting information provided by the location and color cues during the second stage of tracking based on the efficient conflict strategy and excellent DSmT combination model.

CONCLUSION

This study mainly addressed the visual multi-object tracking in the presence of cluttered scenes based on DSmT and the introduced approach has been tested and evaluated. Experimental results have been demonstrated that the introduced approach ameliorated the interference immunity for tracking multi-object. Especially, the tracking accuracy and robustness can be improved while not affecting the real-time characteristics of video image. Therefore, the suggested approach is a useful method for dealing with high conflict between evidences and improving the performance of PF, the approach exhibited a significantly better performance for tracking efficiency and accuracy than a conventional PF and it can easily be generalized to deal with additional cues and objects in the presence of cluttered scenes.

ACKNOWLEDGMENT

The authors express grateful thanks to the support provided by the Natural Science Foundation of Shaanxi Province, China (No.2012JM8004).

REFERENCES

Abdallah, F., A. Gning and P. Bonnifait, 2008. Box particle filtering for nonlinear state estimation using interval analysis. Automatica, 44: 807-808.

Brun, O., V. Teulieve and J.M. Garcia, 2002. Parallel particle filtering. J. Parall. Distrib. Comp., 62: 1186-1202.

Crisan, D. and O. Obanubi, 2012. Particle filters with random resampling times. Stochast. Proc. Appl., 122: 1332-1368.

Dezert, J. and F. Smarandache, 2004. Advances and Applications of DSmT for Information Fusion. American Research Press, Rehoboth.

Djuric, P.M., J.H. Kotecha and J. Zhang, 2003. Particle filtering: A review of the theory and how it can be used for solving problems in wireless communications. IEEE Signal Proc. Mag., 20: 19-38.

Han, Z.J., J.B. Jiao, B.C. Zhang, Q.X. Ye and J.Z. Liu, 2011. Visual object tracking via sample-based adaptive sparse representation. Patt. Recognit., 44: 2170-2183.

Kazuhiro, H., 2009. Adaptive weighting of local classifiers by particle filters for robust tracking. Patt. Recog., 42: 619-628.

Lao, Y., J. Zhu and Y.F. Zheng, 2009. Sequential particle generation for visual tracking. IEEE T. Circ. Syst. Video Technol., 19: 1365-1378.

Lefevre, O.C., 2002. Belief functions combination and conflict management. Inform. Fusion J., 3: 149-162.

Marcelo, G.S., 2005. Improved sequential Monte Carlo filtering for ballistic object tracking. IEEE T. Aerospace Elec. Syst., 41: 1103-1108.

Tehamova, A., J. Dezert and T. Semerdjiev, 2005. Multi-Target Tracking Applications of Dezert-Smarandache Theory. NATO Advanced Study Institute, Albena, Bulgaria, pp: 66-69.

Yang, H.C., K. Chen and J.A. Yang, 2010. Fusion of communication interception information based on DSmT. Comp. Eng. Appl., 46: 129-132.

Zhao, T. and R. Nevatia, 2004. Tracking Multiple Humans in Crowded Environment. CVPR, Washington D.C., pp: 406-413.