Object Contour Tracking Using Multi-feature Fusion based Particle Filter

Xiaofeng Lu[1,3], Li Song[1,2], Songyu Yu[1], Nam Ling[2]

Institute of Image Communication and Information Processing[1]
Shanghai Jiao Tong University, Shanghai, China
Department of Computer Engineering[2]
Santa Clara University, Santa Clara, California, USA
School of communication and information engineering[3]
Shanghai University, Shanghai, China
Email:{luxiaofeng_sjtu, song_li, syyu}@sjtu.edu.cn, {lsong, nling}@scu.edu

***Abstract*—In this paper, a novel object contour tracking framework integrating independent multi-feature fusion object rough location and region-based temporal differencing model is proposed. In our model, the object rough location tracking is realized by color histogram and Harris corner features fusion method in particle filter framework. Thus it can achieve more robust tracking performance in many challenge scenes. And this particle filter framework is based on our previous CamShift guided particle filter [7]. With the rough object location, efficient region-based temporal differencing model is adopted for object contour detection, then this method is faster and more effective compared to active contour models or conventional global temporal differencing models. Moreover, exact contour tracking result can be used to guide the particle propagation of next frame, to enable more efficient particle redistributions and reducing particle degeneration. Experimental results demonstrate that this proposed method is simple but effective in object location and contour tracking.**

Keywords-object tracking; contour tracking; multi-feature fusion; particle filter; harris corner; region-based temporal differencing

I. INTRODUCTION

Interested object tracking in video sequences has always been a very important problem in several research and application fields, such as behavior understanding, video compression, robotics, surveillance, and so on. Tracking accuracy, multi-object tracking, occlusion and disocclusion adaptation are still challenge issues about object tracking [1]. In object tracking, the key task is how to associate object accurate location over continuous frames through different methods. But most methods simplify this task to represent the tracked objects with rectangles or ellipses [5][6][7]. In contrast, exact object contours are represented instead of the rough locations, especially for non-rigid targets [8]. And this exact object location representation is much more useful for further behavior analysis like abnormity detection.

Particle filter is a popular optimal Bayesian algorithm for nonlinear and non-Gaussian object tracking and it obtains the most likely posterior estimation based on sequential Monte Carlo framework. Due to its multiple hypothesis property and its good tracking result, we can apply this method to any state-space model to estimate the trajectory of an target in video sequences effectively [2][3]. Following the early work of Isard [2], a variety of extended particle filter frameworks have been proposed for object tracking tasks [4][5][6]. Reference [7] is our previous work, a CamShift guided particle filter (CAMSGPF). In CAMSGPF, we enhance conventional particle filter algorithm by integrating mean shift algorithm to guide particle propagation to improve the tracking accuracy in illumination changing and severe object deforming scenes.

In recent years, object tracking methods based on a variety of object detection per frame in a video sequence have been proposed. And foreground extraction and background extraction are two main categories. In background modeling category, one of the most widely used approaches is the Gaussian mixtures models (GMM) [9]. However, the Gaussian assumption for every pixel intensity distribution is not always satisfied in dynamic scenes and GMM may fail in many scenes. In [10][11], a local binary pattern (LBP) algorithm and its extended model are proposed for object detection. However, LBP algorithm is not so efficient since its sensitivity to noise and long histograms, then it can not realize real-time detection since its high computational load. At the same time, temporal differencing method (TDM) [12][13] computes the difference between adjacent two or more frames in a sequence to obtain the moving regions as a foreground extraction model. TDM is suitable for real-time object detection because of its fast computation. However, global TDM is easy to detect false object and cause the ghost in dynamic background scenes for background noise or illumination changing.

The particle filter method and background models are always depending on the pixel or region statistic features. And they will be sensitive to illumination, other objects interference, or target deformation. So the boundary-based object tracking models, mostly based on active contour models (ACMs), are widely adopted in vision application. ACMs or snake model was first proposed by Kass et al. [14] for target contour detection. This method utilized an energy minimizing function to a deformable contour so as to approach the true target boundaries. ACMs can be broadly classified into the parametric models and geometric models according to their representation and implementation difference. Parametric model based on region segmentation [15] iteratively searches its new position in current frame to reach object boundaries by minimizing the energy function. However, this model has heavy sensitivity to initial contour settings and will be confused the real objects with similar color regions. Geometric models [16][17] are superior to parametric models since they can automatically handle the topological changes of the propagating curve. These geometric models are motivated by a curve evolution approach, and the deformation is implemented using a level set method [18]. The main advantage of ACMs is that it can detect the deformable object contours accurately. However, they require initial contour settings and need a very long computational time to achieve the energy convergence.

In general, the selected features, that can efficiently distinguish the objects of interest from the background, determine the performance of tracking methods. Regular features like color, motion, texture, and frame difference are modeled by kinds of tracking algorithms [6][7][8]. Recent approaches combine different features for robust object tracking [19] or select the most discriminative features from candidates online [20].

In this paper, we propose a new object contour tracking model framework to integrating independent multi-feature fusion object rough location and region-based temporal differencing model. Our work has three major contributions: 1) the object rough location tracking is realized by color histogram and Harris corner features [23] fusion method in particle filter framework. Thus it can achieve more robust tracking performance in many challenge scenes like illumination changes, transformation, short time occlusion; 2) region-based temporal differencing model (RTDM) is adopted in our object contour detection step, and this effective region is the rough location tracking result. Thus, this model is simpler and faster compared to snake or level set method. In additionally, it has more efficient computing region than conventional global TDM, and it can tolerate complex background or similar color objects to get more robust object contour; 3) contour tracking result, that is exact object boundary location, can be used to guide the current particles' propagation, which enables more effective redistributions of particles towards locations associated with high probability and reduces the particle degeneration possibility. In this case, our new particle filter framework just uses a very small number of particles (N=20 is used).

This paper is organized as follows. Section II gives general description of the algorithm framework. Section III describes the proposed algorithm scheme in detail. And Section IV shows the experimental results. Conclusions are presented in Section V.

II. General Framework Description

The proposed object contour tracking general framework can be divided into two procedures, i.e. the object tracking initial procedure and the exact contour-guided multi-feature fusion object contour tracking procedure as show in Fig. 1.

In initial procedure, we obtain the object rough location with a rectangle manually as a reference template. Based on this rough location, the state vectors initialization of particles is finished, and we can acquire the reference color histogram. Similarly, the object boundary in rough location is computed by Canny operator [21] for Harris corner extraction [23] to form reference Harris corner feature vector. Next, object initial contour is used to guide particle propagation in current frame. Independent color and Harris corner features are integrated by Bayesian filters to produce a detailed representation of the object of interest. When this object rough region is obtained, the model applies the RTDM to candidate location to get exact contour of this tracked object.

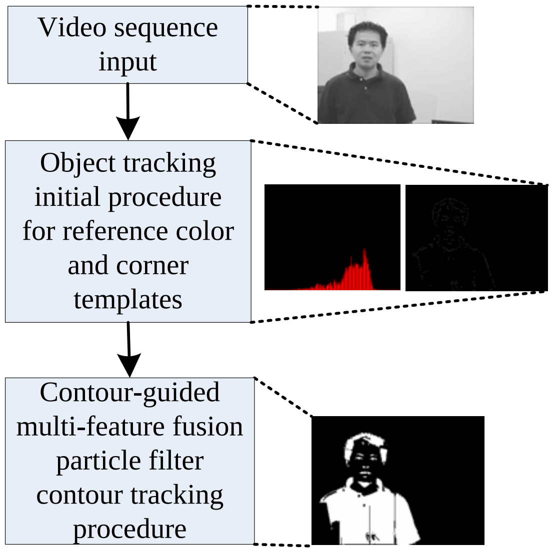

Figure 1. Block diagram of the proposed general algorithm framework.

III. The Proposed Algorithm Procedures

These two procedures about object location initialization and object contour tracking are described in detail as follows.

A. *Object Contour Initial Algorithm Procedure*

Fig. 2 shows the block diagram of the object contour initial procedure. First, the initial frame is captured and we define the tracked interested object with a rectangle manually. Then, we begin to initialize the particle filter state vector based on our previous CAMSGPF framework [7] in that rectangle area.

1) *State vectors of particles initialization:* The candidate is modeled by a state vector as in (1). And we define the rectangle centered at coordinate (x_c, y_c) with width w and height h as the initial rough location. Equation (2) and (3) show the nonlinear state transition model f and observation model E. Then x_k presents the object state and z_k presents the object observation at time k, with n and θ denoting process and observation noise. The posterior density function $p(x_k|z_{1:k})$ is approximated by a set of N samples (particles) with weights $\{x_k^i, w_k^i\}_{i=1}^N$. Given the posterior density function

$p\left(x_{k-1} \middle| z_{1:k-1}\right)$, under the first order Markov assumption, regressive Bayesian theory is optimal.

$$x = \left(x_c, y_c, w, h\right)^T \tag{1}$$

$$x_k = f_k\left(x_{k-1}, n_{k-1}\right) \tag{2}$$

$$z_k = E_k\left(x_k, \theta_k\right) \tag{3}$$

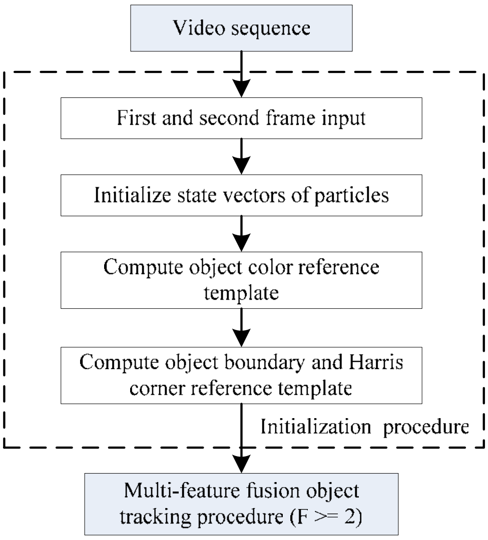

Figure 2. Block diagram of othe object initial algorithm procedure.

In particle filter framework, each sample is weighted in terms of the observation with probability computing formula (4). In this method, the proposal importance sampling distribution function $q\left(x_k^i \middle| x_{0:k-1}^i, z_k\right)$ is substituted by prior distribution $p\left(x_k^i \middle| x_{0:k-1}^i\right)$. And the likelihood distribution is presented as (5). Here $D[]$ is Bhattacharyya similarity metric [24] to compare reference template h_0 and candidate template $h[x_k]$.

$$w_k^i \propto [p\left(z_k \middle| x_k^i\right) p\left(x_k^i \middle| x_{0:k-1}^i\right)] / q\left(x_k^i \middle| x_{0:k-1}^i, z_k\right) \tag{4}$$

$$p\left(z_k \middle| x_k^i\right) \propto \exp(-\lambda D^2\left[h_0, h\left(x_k^i\right)\right]) \tag{5}$$

2) Harris corner computing: First, we utilize a temporal differencing [13] between the adjacent two frames to obtain the subtracted result as Fig.3 (c) shown. However, the differencing boundary is not accurate enough. To solve this problem, we adopt Canny edge detector for current frame, and apply a logic AND operation with differencing boundary image to obtain the final motion contour image as Fig. 3(d)(e).

In order to reduce the computational cost for measuring continued contours, we choose a set of Harris feature points to represent the object contour that can keep significant object contour information. Harris corner detector [23] is applied to obtain Harris corner feature, and that is insensitive to illumination changing, rotation or zoom operation. Generally, Harris operator is realized by calculating each pixel's gradient and judge the absolute gradient values in two directions. Harris corner detection algorithm is defined as follows:

$$M = G(\boldsymbol{\sigma}) \otimes \begin{bmatrix} g^2_x & g_x g_y \\ g_x g_y & g^2_y \end{bmatrix} \tag{6}$$

$$R = \det(M) - k \cdot \mathrm{Tr}^2(M),\ 0.04 \le k \le 0.06\,. \tag{7}$$

Where $G(\boldsymbol{\sigma})$ is a Gaussian function with deviation $\boldsymbol{\sigma}$; g_x and g_y is the first derivatives of pixel (x, y). $\det(M)$ means the determinant of matrix M, and Tr denotes the trace of matrix M. And R is the corner strength function at one pixel. If the strength is above a particular threshold, then a Harris corner is found. The Harris corner is presented as Fig. 3(f).

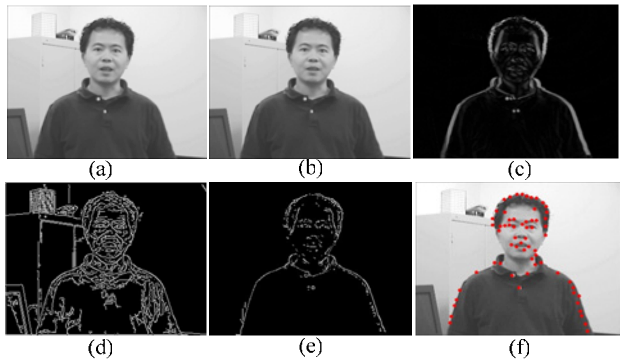

Figure 3. Object tracking initial procedure: (a) and (b) are adjacent frames; (c) result of temporal differencing model; (d) result of Canny edge detection; (e) result of AND logic operation; (f) result of Harris corner detection.

B. Object Contoure Tracking Algorithm Procedure

1) Partilce propagation: When we obtain the initial reference color histogram template and reference Harris corner set, the system begins to getting into the object contour tracking algorithm procedure as Fig. 5 shown. First, the dynamic state transition model is defined as (8). This propagation mechanism satisfied to a standard 2nd order autoregressive process is an efficient particles' evolution.

$$\hat{x}_k = x_{k-1} + \frac{1}{n}\sum_{l=1}^{n}\left|x_{k-l} - x_{k-(l+1)}\right| + n_k \tag{8}$$

2) Color weight estimation: After particle propagation, we can get N candidate particle samples. In our previous work CAMSGPF, we employ Bhattacharyya distance and Bhattacharyya coefficient [24] to present the similarity measurement between one candidate and the reference color template as in (9) shown, where p_{ref} and q_{can} are two histograms. From this similarity measurement, we can use (10) to define the color weight of this candidate. And $\boldsymbol{\sigma}$ is the variance determined empirically.

$$d_{color} = \sqrt{1 - \sum_{i=1}^{N} \sqrt{p_{ref}(y_o^i) q_{can}(y_t^i)}} \quad (9)$$

$$weight_{color}^{j} = (1/\sqrt{2\pi\sigma}) \exp(-d_{color}^{2} / 2\sigma^{2}) \quad (10)$$

3) Harris corner weight estimation: For every object candidate, we must compare the similarity bewteen candidate Harris corner set and reference Harris corner set. In this procedure, we use similarity Hausdorff distance [22] to represent this distance between two point sets. The Harris corners are changing their positions and number, and simularity Hausdorff distance does not need to establish the point-to-point relationships. To two finite point sets, A={ $a_1, a_2, ..., a_n$ } and B{ $b_1, b_2, ..., b_n$ }, the Hausdorff distance is defined with (11) , (12) and (13) [22]. Where $d(A,B)$ is the Hausdorff distance between set A and set B, and $h(A,B)$ is the distance from set A to B. $\|\bullet\|$ defines the Euclidean norm between two corner set,

$$d(A,B) = \max[h(A,B), h(B,A)] \quad (11)$$

$$h(A,B) = \max_{a \in A} \min_{b \in B} \|a_i - b_j\| \quad (12)$$

$$h(B,A) = \max_{b \in B} \min_{a \in A} \|b_i - a_j\| \quad (13)$$

Simultaneously, we can define the Harris corner weight of the candidate from the similarity measurement based on Hausdorff distance as (14) shown comparing to color weight.

$$weight_{contour}^{j} = (1/\sqrt{2\pi\sigma}) \exp(-d_{contour}^{2} / 2\sigma^{2}) \quad (14)$$

$$weight_{t}^{j} = \alpha weight_{t\,(color)}^{j} + \beta weight_{t\,(contour)}^{j} \quad (15)$$

4) Object rough region estimation: From (15), we can obtain the feature weight of this current candidate. α and β mean constant factors. This multi-feature fusion scheme combines color and corner feature with different effective feature similarity measurements, and it achieves object location that can be tolerant to illumination and object deformation. Through weight sorting of all particles, it is easy to get the object rough location region in sequence.

5) Region-based TDM: As Fig. 4 shown, here we apply a region-based TDM algorithm to this rough object location to obtain the object binary region image. And the exact object contour can be received followed by morphology operation and contour finder. Generally, the region size we need to compute is much smaller than the whole image, So this fast algorithm can really reduce the computational cost of whole algorithm structure.

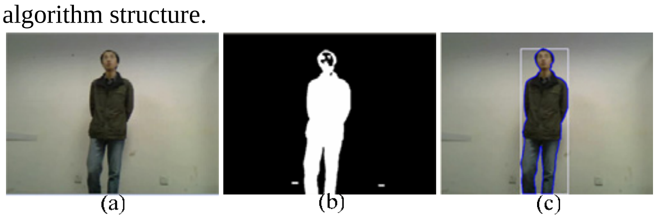

Figure 4. Region-based TDM results: (a) current frame. (b) object binary region image. (c) object rough location and exact contour.

6) Object exact location guided particle propagation: If the random noise in particle propagation is small, the particle resample impoverishment will be obvious. On the contray, if the random noise is large, it will need more particles to realize exact state prediction [6].

$$n_k = r_k * n'_k = (a * wr_k + b * hr_k) * n'_k \quad (16)$$

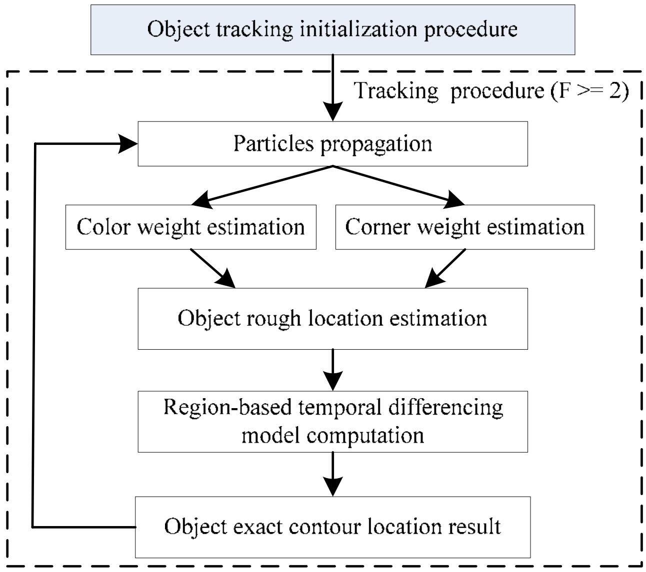

Figure 5. Object contour tracking algorithm procedure.

Equation (16) is the new contour scale guided particle propagation noise model we proposed, where n_k is new noise vector, r_k is exact contour scale coefficient, and n'_k meets zero-mean Gaussian distribution. In this model, we define a and b as vertical and horizontal coordinate noise coefficient empirically, and they satisfy the definition $a + b = 1$ ($a = 0.7$, $b = 0.3$). Then the exact contour object location trends can be used to guide the current particles' propagation, which enables a more effective redistributions of particles towards locations associated with high probability and reduce the particle degeneration possibility. In this case, our new particle filter framework just uses a very small number of particles (N=20 is used) to realize exact object rough location and exact contour tracking.

IV. EXPERIMENTANL RESULTS

In this section, to verify the performance of this proposed method, some common test video sequences from Toronto University and our own datasets are used. And the conventional particle filter (CPF) [3] and CAMSGPF [7] are also performed for comparison.

A. *Multi-feature Fusion Rough Object Location Results*

Fig. 6 compares the object rough face location tracking results presented by CPF [3], CAMSGPF [7] and our proposed method. In this David2_WMM_xvid video sequence, illumination on object face is changing rapidly, and object scale is changing when David moving around the camera. We can see that CPF will lose tracked object quickly, and CAMSPF can not deal with illumination changing very well. While our proposed method can tolerate object scale and illumination changes during the whole tracking process by fusion color and corner features.

Another advantage of our proposed multi-feature fusion tracking strategy is the tracking robustness to the short occlusion, i.e. our proposed model can quickly re-obtain satisfactory tracking results even after a short whole object occlusion as Fig. 7 shown. One can observe that CPF method will quickly lose the tracked object when same color feature object appears in frame 186. On the other hand, when a short occlusion happens, CAMSGPF method can not recover the target and turn into tracking error. But to our proposed method, short occlusion will not result in loss of track, and it can retarget the object exactly after occlusion.

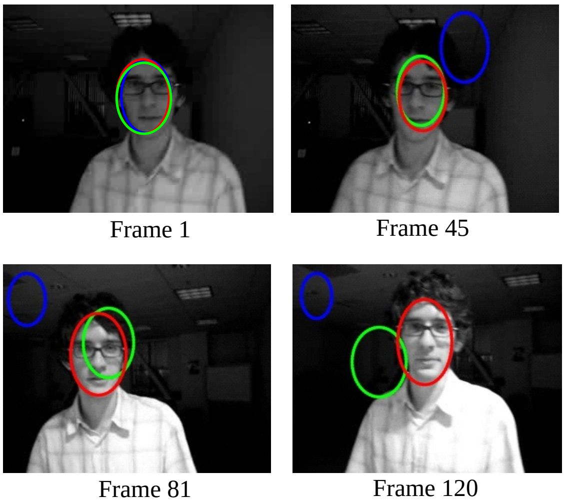

Figure 6. David2_WMM_xvid video sequence: Blue ellipse is the object tracking result from CPF; Green ellipse is the tracking result from CAMSGPF; And Red ellipse is the tracking result of our proposed.

Figure 7. Anvar1GlobalNormal video sequence: Blue ellipse is the object tracking result from CPF; Green ellipse is the tracking result from CAMSPF; And Red ellipse is the tracking result of our proposed.

B. *Contour Tracking Results*

With multi-feature fusion particle importance evaluation scheme, we can obtain more exact rough object location than conventional particle filter. Then, based on this rough location, we further assess the algorithm performance on object exact contour extraction model. In this proposed algorithm, we apply the exact region-based temporal differencing model to extract object contour as Section III part B described. This model is simpler and faster compared to snake or level set method. In additionally, it is more efficient than conventional global TDM to get more robust object contour. Fig. 8 presents the Lu_1 video sequence tracking result. We initialize the tracking object with a green rectangle in frame 1, and this method is capable to get object rough location with green rectangle and exact contour with blue connection line accurately as frame 65, frame 135, frame 228 shown. And in another experiment for Zhang_2 video sequence, the illumination is always changing and the tracking results are presented in Fig. 9. One can observe that in this sequence, this method can tolerate object deformation, but the contour in face region is not very good for the severe illumination affect to temporal differencing model.

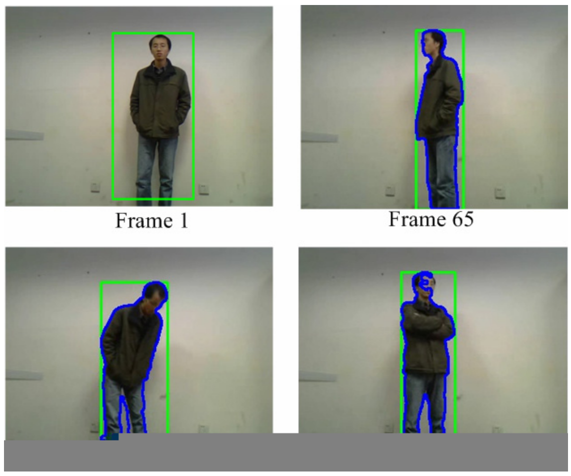

Figure 8. Object rough location and exact contour tracking in room scene for Lu_1 video sequence.

Figure 9. Object rough location and exact contour tracking in weak illumination scene for Zhang_2 video sequence.

V. CONCLUSION

In this paper, a new object contour tracking framework is proposed. This method combines multi-feature fusion strategy based object rough location and exact region-based object contour extraction for accurate and robust object contour tracking. Specifically, in our proposed model, the object rough location is realized by color histogram and Harris corner features fusion method in particle filter framework. Region-based TDM is applied for exact contour detection, which is simpler than active contour models. And in turn, object exact contour scale is also applied to guide new particle propagation procedure. The experimental results demonstrate the improved performance of this proposed model. In the future work, we are planning to embed accurate object local motion information into contour exact step to improve the contour edge accuracy based on TDM. And it is also need to realize traditional active contour model for algorithm comparison.

ACKNOWLEDGMENT

This research was supported in part by NSFC (60702044), NSFC (60902073), 973 Program (2010CB731400) and the 111 project.

REFERENCES

[1] Alper Yilmaz, Omar Javed, and Mubarak Shah, "Object tracking: a survey," ACM Computing Surveys, Vol.38, No.4, Article 13, Dec. 2006.

[2] Isard M., Blake A., "CONDENSATION-conditional density propagation for visual tracking," International Journal of Computer Vison, 29(1), pp.5-28, 1998.

[3] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," Statistics and Computing, 10(3), pp.197-208, 2000.

[4] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters fot online noninear/non-Gaussian Bayesian tracking," IEEE Trans. on Signal Processing, Vol.50(2), pp.174-188, 2002.

[5] Zulfiqar Hasan Khan, Irene Yu-Hua Gu, and Andrew G. Backhouse, "A robust particle filter-based method for tracking single visual object through complex scenes using dynamical object shape and appearance similarity," J Sign Process Syst, Vol.65, pp.63-79, 2011.

[6] S.K. Zhou, R. Chellappa and B. Moghaddam, "Visual tracking and recognition using appearance-adaptive models in particle filter," IEEE Trans. on Image Processing, Vol.13, No.11, Nov. 2004.

[7] Z. Wang, X. Yang, Y. Xu, and S. Yu, "CamShift guided particle filter for visual tracking," ACM Pattern Recognition Letters, Vol.30, pp. 407-413, Mar. 2009.

[8] Ling Cai, Lei He, Yamasita Takayoshi, Yiren Xu, Yuming Zhao, and Xin Yang, "Robust contour tracking by combining region and boundary information," IEEE Trans. on Circuits and System for Video Technology, in press.

[9] C. Stauffer and W.E.L. Grimsom, "Adaptive background mixture models for real-time tracking," Proc. IEEE CS Conf. Computer Vision and Pattern Recognition, Vol.2, pp.246-252, 1999.

[10] M. Heikkila and M. Pietikainen, "A texture-based method for modelling the background and detecting moving objects," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.28, No.4, pp.657-662, 2006.

[11] Xue Gengjian, Sun Jun, and Song Li, "Dynamic background subtraction based on spatial extended center-symmetric local binary pattern," IEEE International Conference on Multimedia & Expo, ICME, pp.1050-1054, July, 2010.

[12] Q. Cai and J. K. Agarwal, "Tracking human motion in structured environments using a distributed-camera system," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.21, No.11, pp.1241-1247, Nov. 1999.

[13] A. J. Lipton, H. Fujiyoshi, and R. S. Patil, "Moving target classification and tracking from real-time video," in Proc. of the IEEE Workshop on Applications of Computer Vision, pp.8-14, Oct. 1998.

[14] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: active contour models," International Journal of Computer Vision, Vol.1(4), pp.321-331, 1988.

[15] C. Xu and J. L. Prince, "Snakes shapes, and gradient vector flow," IEEE Trans. on Image Processing, Vol.7(3), March, 1998.

[16] N. Paragios and R. Deriche, "Geodesic active contours and level sets for the detection and tracking of moving objects," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol.22(3), pp.266-280, 2000.

[17] R. Goldenberg and R. Kimmel, "Fast geodesic active contours," IEEE Trans. Image Processing, 10(10), pp.1467-1475, 2001

[18] J. A. Sethian, Level set methods, Cambrdge Uni. Press, 1996.

[19] V. Takala and M. Pietikinen, "Multi-object tracking using color, texture and motion," in Proc. IEEE Conf. Computer Vison and Pattern Recognition, pp.1-7, 2007.

[20] R. Collins, Y. Liu, and M. Leordeanu, "On-line selection of discriminative tracking features," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol.27, No.10, pp.1631-1643, 2005.

[21] J. Canny, "A computational approach to edge detection," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.8, pp.679-698, 1986.

[22] D. P. Huttenlocher, G. A. Klanderman, and W. A. Rucklidge, "Comparing images using the Hausdorff distance," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.15, pp.850-863, 1993.

[23] C. Harris and M. J. Stephens, "A combined corner and edge detector," In Alvey Vision Conference, pp.147-152, 1988.

[24] T. Zhang, S. Fei, X. Li and H. Lu, "An improved particle filter for tracking color object," IEEE Computer Society International Conf. on Intelligent Computation Technology and Automation, pp. 109-113, 2008.