



Title	Particle Filter for Targets Tracking with Motion Model
Author(s)	Pang, GKH; Choy, KL
Citation	The 8th International Conference on Industrial and Information Systems (ICIIS), Peradeniya, Sri Lanka, 17-20 December 2013. In International Conference on Industrial and Information Systems, 2013, p. 128-132, article no. 6731968
Issued Date	2013
URL	http://hdl.handle.net/10722/203994
Rights	International Conference on Industrial and Information Systems. Copyright © I E E E.

PARTICLE FILTER FOR TARGETS TRACKING WITH MOTION MODEL

Grantham K.H. Pang
Department of Electrical and Electronic
Engineering,
The University of Hong Kong,
gpang@eee.hku.hk

K. L Choy
Department of Industrial and Systems Engg.
The Hong Kong Polytechnic University
Hung Hom, Kowloon, Hong Kong
kl.choy@polyu.edu.hk

ABSTRACT

Real-time robust tracking for multiple non-rigid objects is a challenging task in computer vision research. In recent years, stochastic sampling based particle filter has been widely used to describe the complicated target features of image sequence. In this paper, non-parametric density estimation and particle filter techniques are employed to model the background and track the object. Color feature and motion model of the target are extracted and used as key features in the tracking step, in order to adapt to multiple variations in the scene, such as background clutters, object's scale change and partial overlap of different targets. The paper also presents the experimental result on the robustness and effectiveness of the proposed method in a number of outdoor and indoor visual surveillance scenes.

Index Terms— Target tracking, particle filter, kernel density estimation.

1. INTRODUCTION

Target tracking is a key technique which is widely used in visual surveillance system. This research topic has attracted much attention in the computer vision research field. The core issue for target tracking is to localize the target and label their identity. We consider the solution to multiple-target tracking problem as the following steps. First, we reconstruct the background scene (with the Kernel Density Estimation technique) and extract the foreground binary image for the moving targets. Secondly, we localize the multiple targets through the binary image. Thirdly, we identify the target and compute their similarity with particle filter model, in order to achieve the goal of labeling the targets. Color histogram information of target region is employed at this stage to determine the similarity of different objects.

Our target tracking research focuses on the multiple target tracking in the video surveillance applications. We initialize the video clip with a construction of the background scene and then to extract the foreground moving

targets. Subsequently, we employ the particle filter tracking model for targets tracking. The core problem is to recognize and identify the target with the scene clutters (target occlusion, scale variance of object).

2. RELATED WORK

A five-dimensional space model was proposed in [1], which consisted of two components for the optical flow and three for the intensity in the color space. Clutters like tree leaves swinging and ocean's wave in the background scene were suppressed. It defined a multivariate fixed bandwidth kernel estimator. The approach in [1] utilized optical flow as two of the components in a 5-dimensional feature space, which was computationally expensive and not suitable for real-time processing.

The particle filter theory, which is also known as condensation algorithm, was first used by Isard and Blake in 1997 [2] to solve the target tracking problem in computer vision. Particle filter offers a degree of robustness to unpredictable motion and can correctly handle complicated nonlinear measurement models. The idea was originated earlier in 1993 by N.Gordon and D.Salmond [3] with the analysis of Bayesian state estimation using particle filter theory, which also compared with the traditional Kalman filter and the Extended Kalman filter to show particle filter's advantage when facing the nonlinear/non-Gaussian statistical model. It can be made more accurate than either Kalman filter or Extended Kalman filter because of particle filter's sufficient sampling points and so it approaches the Bayesian optimal estimate.

Z. Khan and T. Balch [4] proposed a Markov Chain Monte Carlo Computing (MCMC) based joint particle filter model for multiple interacting targets, and later they also used this model for the experiment of multiple ants tracking with known and unknown target number. Their publications gave explicit test result of interacting process of the targets. C. Yang [5] used modified hierarchical particle filter to track the moving target, with the information of color and edge orientation histogram features, which demonstrated a good result when facing short-time occlusion.

In [6], the particle filtering technique was used for position and velocity tracking of a target in a radar system. The authors have also shown that the technique is better in accuracy when compared with the Kalman filtering technique. A multi-aspect target tracking problem was

presented in [7], where a continuous-valued affine model was incorporated in the state vector for multi-aspect modeling. Three particle filtering algorithms were compared, and an algorithm that involved two likelihood functions and a re-weighting scheme was the best as it balanced the diversity and focus of the particles.

Yuan Li et al. [8] proposed a cascade particle filter with discriminative observers to track fast moving target, and abandoned the traditional assumption of motion continuity of the target. Kenji Okuma and David Lowe [9] proposed the “boosted particle filter” tracking model, which provokes the Adaboost method to construct the proposal distribution model that can improve the effectiveness of detecting objects leaving and entering the scene.

Recent advance in target tracking research has shown that an increasing attention was paid to segmenting the different targets when they appear overlapped with each other. For example, Bose and Wang [10] divided the states into three labels: group, object and fragment, using an online generic tracking model without previous model learning. J. Sullivan and S. Carlsson [11] also employed the “group object” model to describe the target interaction, which is applied to the case of football players tracking. Therefore, the particle filter method can be incorporated with other novel tracking clues to achieve a robust multiple tracking results.

Zhang et al. [12] developed an SVD based Kalman Particle Filter for visual tracking. The sigma samples are generated by SVD of an eigen-covariance matrix, which overcomes the ill-conditioning issue when solving the square root of a covariance matrix in a typical unscented Kalman filter used for target tracking. Robust performance was obtained in a comparison among the standard particle filter, unscented particle filter and the developed method. In [13], the blob image of the object is treated as a matrix and singular values are adopted to construct the feature model. In the particle filter scheme, the traditional resampling step is substituted for Metropolis-Hastings sampling. The proposed method has claimed that it can perform good tracking even when there are confusing objects in the scene. The weight variance and number of particles in the method are less when compared with the conventional method.

Hossain and Lee [14] focused on developing a new observation model for object tracking in the challenging situation of occlusion. The number of particles could vary in the process of optimization. In [15], the paper was focused on vehicle counting using particle filter. The heuristic of motion coherence was used to help locate and group particles in video sequences. Later, color histogram was also used in the tracking of objects.

Recently, some papers are focused on the computational issues related to the implementation of the particle filter algorithm using high performance processors. Cho et al. [16] designed the circuit to implement the algorithm using VHDL (VHSIC Hardware Description Language) in an FPGA (Field Programmable Gate Array). All the functions of the particle filter tracking system are implemented in an FPGA, and real time objects tracking has been with good performance.

In [17], a particle filter algorithm has been implemented using SIMD (Single Instruction, Multiple Data) processors for the tracking of non-rigid objects. A color histogram of the target is provided to the tracker initially, and the algorithm would search each subsequent frame for the new location of the target using Bayesian estimation. The method is suitable for tracking non-rigid objects since the color histogram is relatively independent of the target deformation. The method is also robust to occlusion and variations in the color of the background. The high computational cost of the method has been met by the use of an SIMD linear processor array. Sutharsan et al. [18] also proposed an SIMD particle filter for multi-target tracking. By balancing the load (i.e. the number of particles) processed by each processing element, the algorithm has been optimized for minimum computational time. Also, a distributed resampling method which requires exchange of fewer particles among processors has been used.

In this paper, the objective is to develop a robust particle filter based target detection and tracking method using the color histogram. We have focused on target motion model to help with object matching between different frames. The experimental results have shown the robustness of our method in challenging cases, such as the tracking of scaling target, and when targets are overlapped caused by interactions among different objects.

3. BASIC THEORY AND OUR METHOD

In this section, we present the basic theory involved with our target detection and tracking. The implementation is based on the theoretical foundation in this section.

3.1. Kernel Density Estimation Based Background Model

The kernel density estimation (KDE) modeling for background is used in this paper. The main objective is to capture the fast changing of local background scenes. We can take a certain kernel density function to determine the observation pixel probability of belonging to background scene. This estimation is based on the chosen historic consecutive pixels' intensity. The background scene can then be estimated from the recent historic samples incorporated with the empirical threshold value.

Theoretically, the KDE method is a kind of smoothing method to reconstruct the probability density function for a series of samples. Although density estimators such as the histogram density estimator can be made to be asymptotically consistent, others are often either discontinuous or would converge at slower rates than the kernel density estimator. Rather than grouping observations together in bins, the kernel density estimator can be thought to place small “bumps” at each observation, determined by the kernel function.

3.2. Particle Filter Based Tracking

The primary goal of “particle filter” is to determine the posterior distribution $P(x_t | Z_t)$ over the current joint

configuration of the target X_t at the current time step t , given all the observations $\{Z_1, \dots, Z_t\}$ up to that time step.

The aim of employing the filter is to apply Bayesian theorem at each time-step, obtaining a posterior $p(x_t | Z_t)$ based on all available information.

$$p_t(x_t | Z_t) = \frac{p_t(Z_t | x_t) p_{t-1}(x_t | Z_{t-1})}{p_t(Z_t)} \quad (1)$$

A model for the expected motion between time-step is adopted. According the conditional distribution property, the above function can be written as:

$$p_t(x_t | Z_t) = \frac{p_t(Z_t | x_t) p_t(x_t | x_{t-1}) p_{t-1}(x_{t-1} | Z_{t-1})}{p_t(Z_t)} \quad (2)$$

Bayesian theory assumes that $p_t(Z_t | x_t)$ is Gaussian, "particle filter theory" is based on the Gaussian distribution to generate K candidate regions in the next frame. This method determines the above formula in the following steps:

$$p_{t-1}(x_{t-1} | Z_{t-1}) \mapsto p_{t-1}(x_t | Z_{t-1}) \mapsto p_t(x_t | Z_t)$$

- 1) Multiply by dynamic transition $P(x_t | x_{t-1})$;
- 2) Multiply by observation density $P(Z_t | x_t)$.

The simulation uses the idea of a particle set: "particle" means a list of n pairs $(x_i, \pi_i), i = 1, \dots, n$ and $\pi_i \in [0, 1]$ with $\sum_{i=1}^n \pi_i = 1$. Every particle set represents a probability distribution $p(x)$.

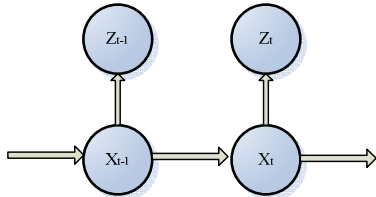


Figure 1. Hidden states (eg. Object location, scale)

We perform forward inference using the Bayesian filtering distribution:

$$p_t(x_t | Z_t) = \alpha \times p_t(Z_t | x_t) p_t(x_t | x_{t-1}) p_{t-1}(x_{t-1} | Z_{t-1}) \quad (3)$$

The four terms represent the corresponding four states above. They are current object state, observation model, transition model, and previous object state respectively.

-- Prior Distribution: $P(x_0)$

- Describes initial distribution of object states

The methods to determine the prior distribution could be based on an object detector, or based on the user input.

-- Transition Model: $P(x_t | x_{t-1})$

- Specifies how objects move between frames

The method to determine it:

- The next state estimation is based on a Gaussian window around the current state. We take into account the velocity information of previous states.

A velocity based motion model can be used as follow:

$$\hat{X}_t = A \hat{X}_{t-1} + w_t \quad (4)$$

This predicts the position of re-sampling candidate regions based on the previous state and velocity as well as noise w_t , here w_t is a normal distribution stochastic noise.

-- Observation Model: $P(Z_t | x_t)$

- It specifies the likelihood of an object being in a specific state (i.e. at a specific location)

3.3. The Proposed Method

Our developed particle filter tracking model is based on object motion model and color histogram as well as using an elliptical tracker.

The particle filter is applied in a color-based context. K candidate regions L_1, L_2, \dots, L_k are generated at the initialization stage. In our experiment, K equals to 200~500 and that depends on different region size. The generated sampling regions are based on the normal distribution centering on region at time t :

$$u_x = x_0, \sigma_x < Width / 3.92; u_y = y_0, \sigma_y < Height / 3.92.$$

This is to ensure that 95% of generated central points of candidate regions are inside the original object region.

Here, x_0, y_0 represent the central point coordinate of previous object region. $Width, Height$ represent the size of previous object region. The elliptical candidate regions are created with central point (x, y) and the horizontal and vertical axes are previous object region $Width$ and $Height$.

Color distributions are used as target models as they achieve robustness against non-rigidity, rotation and partial occlusion. Suppose that the distributions are discretized into m -bins. The histograms are produced with h_t^i that assigns the color at location L_i to the corresponding bin. In our experiments, the histograms are typically calculated in the RGB space using $8^3 = 256$ different colors.

The similarity of candidate region L_i at t with object region at time $(t-1)$ is based on a distance metric D_i between histograms h_{t-1} and h_t^i of the i th candidate region at time t , $\beta > 0$ is the direct proportional coefficient:

$$p^{(i)}(Z_t | x_t) = \beta e^{-\lambda D_i^2} \quad (5)$$

To weight the sample set, the Bhattacharyya coefficient has to be computed between the target histogram and the histogram of the hypotheses. The distance metric D is as:

$$D_i = \left\| h_{t-1} - h_t^i \right\|_1 \quad (6)$$

h_{t-1}, h_t^i are the color histogram of object region in last frame $t-1$ and the i th candidate region in this frame t .

$p^{(i)}(Z_t | x_t)$ represents the similarity of the color histogram of candidate region and target region in the previous frame. The region centre is the origin of the elliptic region. It is desirable to obtain the samples whose color distributions are similar to the target model. Therefore, small

Bhattacharyya distances correspond to large weights should be considered.

Samples with a high weight (greater than $th1$) will be chosen into the set S_H for computing the estimation position (coordinate of ellipse center) of the object, while others with relatively low weights should not be chosen. Then the estimation coordinate of object position at time t is as follow:

$$\begin{cases} x_t = \sum_{i \in S_H} p^{(i)}(Z_t | x_t) x_t^{(i)} / \sum_{i \in S_H} p^{(i)}(Z_t | x_t) \\ y_t = \sum_{i \in S_H} p^{(i)}(Z_t | x_t) y_t^{(i)} / \sum_{i \in S_H} p^{(i)}(Z_t | x_t) \end{cases} \quad (7)$$

Samples with a low similarity value (less than $th2$) will be eliminated. When the existing samples are less than 40% of total generated ones, the re-generation of K particle samples will be carried out based on the following motion model and the similarity value of existing samples. To resume the K sample regions, the residual samples increase by N times, and $N = \frac{K}{\text{Residual}}$.

Each sample of the distribution represents an ellipse and is given as: $s = \{x, \dot{x}, y, \dot{y}, H_x, H_y\}$. In the expression, x, y specify the location of the ellipse center, \dot{x}, \dot{y} are the motion speed, H_x, H_y are the length of the half axes. We consider a whole sample set, and the tracker handles multiple hypotheses simultaneously.

The sample set is propagated and re-sampling through the application of a first-order dynamic model $s_t = A s_{t-1} + w_{t-1}$ where w_{t-1} is the additive normal distribution noise. Meanwhile A defines the deterministic component of the model. A is a 6×6 matrix that defines the transition model. The motion speed is estimated based on previous M frames (M ranges between 10~20 in the experiment). After the above re-sampling stage, there would be K elliptical sample regions again and the object matching continues in the next frame.

4. EXPERIMENTAL RESULT

A background reconstruction algorithm has been used for the football sequence. The background image is reconstructed using fifty recent historical frames. The background result can be seen below:



Figure 2. The extracted background scene(left) from a $N=50$ original frames starts with the right one.



Figure 3. Foreground blobs connect with the neighbor target

Our method uses both the binary blob and the color histogram feature to track the interactive target. We first initialize the different targets with their foreground blob binary images, and then when different targets get near, use their color histogram feature to identify different targets. The following is the tracking result:

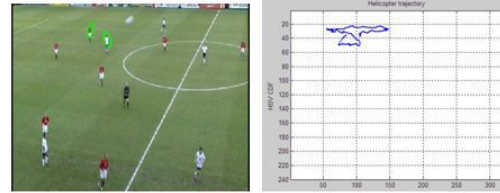


Figure 4. Tracking result and the moving trajectory of one target.

The step of tracking the moving targets makes use of the particle filter based on color histogram information. On football player tracking, both single target tracking and the multiple targets tracking are carried out. The result shows that with the partial occluded and the fast moving scene the algorithm can successfully track the correct targets. The target moving trajectory is also given.



Figure 5. Tracking result when target scale becomes smaller

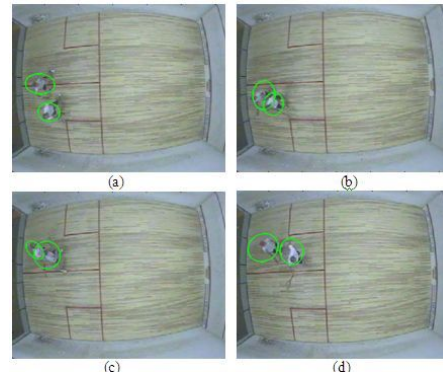


Figure 6. Tracking the interactive targets in indoor squash court

From Fig.5, we have tested the robustness of the tracking method when the target's size varies, as the player dwindles due to his moving away from the camera. The result has shown that our tracker (green ellipse) can adjust its scale to suit the alteration of the target. The detected result shown in Fig.6 illustrates the performance under the partial overlap of close targets in an indoor squash court video. When the two players get closer, part of their detected blob is overlapped by another one, our tracker can still keep the original player when this partial occlusion occurred.

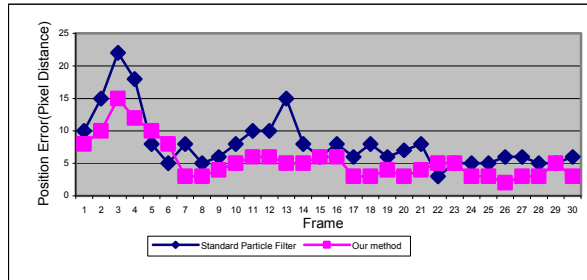


Figure 7. Compare our method with standard particle filter tracking (both have 400 sample particles) Vertical axis indicates the position error of tracker (central point distance compared with ground truth). Horizontal axis indicates 30 chosen frames from the football match image sequence.

Figure 7 shows a comparison between the standard particle filter tracking [2] and our proposed method (with 400 sample particles). Our method shows a better performance in terms of error rate and target scale approximation.

5. CONCLUSION

A robust object tracking method is proposed in this paper. Our method has employed a large number of sampling points (400 sample particles) when applying the particle filter theory to represent the target. Furthermore, the color and velocity along with the RGB color histogram of target region are utilized to determine the similarity of neighboring region in order to find the region with highest likelihood to the original region of target. Through this way, the target's current position in present frame is estimated to support the whole object tracking task. The given result shows the robustness of our method in the challenging cases of target's partial overlap, and target's scaling, both from the outdoor and indoor circumstances. In the future, the performance of our developed method will be compared with some other recent methods reported in our literature survey.

References

[1] A. Mittal and N. Paragios, "Motion-based Background Subtraction Using Adaptive Kernel Density Estimation," Proc. IEEE Conf. Computer Vision and Pattern Recognition, 2004.
 [2] M. Isard and A. Blake, "Contour Tracking by Stochastic Propagation of Conditional Density," Proc. European Conf. Computer Vision, pp. 343-356, 1996.
 [3] N. Gordon, D. Salmond, and A. Smith, "Novel Approach to Nonlinear/Non-Gaussian Bayesian State Estimation," IEE Proc. vol. 140, no. 2, pp. 107-113, 1993.
 [4] Z. Khan, T. Balch and F.Dellaert, "MCMC-based particle filtering for tracking a variable number of interacting targets," IEEE PAMI, Vol.27, No.11, 2005.

[5] C. Yang, R. Duraiswami and L. Davis, "Fast Multiple Object Tracking via a Hierarchical Particle Filter," ICCV 2005.
 [6] S.S. Kambhampati, K.V. Tangirala, K.R. Namuduri and S. Jayaweera, "Particle Filtering For Target Tracking", International Symposium on Wireless Personal Multimedia Communication, WPMC, Sept 12-15, 2004, Italy.
 [7] L. Tang, V.B. Venkataraman and G. L. Fan, "Multi-aspect Target Tracking in Image Sequences Using Particle Filters", Advances in Visual Computing, Lecture Notes in Computer Science Volume 3804, pp 510-518, 2005.
 [8] Y. Li, and H. Ai, "Tracking in low frame rate video: a cascade particle filter with discriminative observers of different life spans," IEEE Trans. on PAMI, Vol.30, No.10, 2008.
 [9] K. Okuma, A. Taleghani, N. Freitas, J. Little, and G. Lowe, "A Boosted Particle Filter: Multitarget Detection and Tracking," ECCV 2004.
 [10] B. Bose, X. Wang, and E.Grimson, "Multi-class object tracking algorithm that handles fragmentation and grouping," In Proc. IEEE Conference on Computer Vision and Pattern Recognition, 2007.
 [11] J.Sullivan and S.Carlsson, "Tracking and labeling of interacting multiple targets," In Proc. ECCV, 2006.
 [12] X. Zhang, W. Hu, Z. Zhao, Y.G. Wang, X. Li and Q.Wei, "SVD Based Kalman Particle Filter for Robust Visual Tracking", Proc. 19th International Conference on Pattern Recognition, pp. 1-4, 2008.
 [13] X. Luo and Y. Huang, "Visual tracking with singular value particle filter", Proc. IEEE International Workshop on Machine Learning for Signal Processing (MLSP 2010), pp. 202-207, August 29 – September 1, 2010, Kittilä, Finland.
 [14] K. Hossain and C.W. Lee, "Visual Object Tracking Using Particle Filter", 2012 9th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI), Daejeon, Korea, November 26-29, 2012.
 [15] C. Bouvie, J. Scharcanski, P.Barcellos and F. Lopes Escouto, "Tracking and Counting Vehicles in Traffic Video Sequences Using Particle Filtering", IEEE International Conference on Instrumentation and Measurement Technology (I2MTC), Page(s): 812 - 815, 2013.
 [16] J. U. Cho, S.H. Jin, X.D. Pham and J.W. Jeon, "Multiple Objects Tracking Circuit using Particle Filters with Multiple Features", 2007 IEEE International Conference on Robotics and Automation, Roma, Italy, 10-14 April 2007.
 [17] H. Medeiros, J. Park and A. Kak, "A Parallel Color-Based Particle Filter for Object Tracking", Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp. 1-8, 2008.
 [18] S. Sutharsan, A. Sinha, T. Kirubarajan, and M. Farooq, "An optimization based parallel particle filter for multitarget tracking", Proceedings of the SPIE, volume 5913, pages 87-98, Jan. 2005.