

# Tracking Object Using Particle Filter and DCT Features

Cong Lin, Chi-Man Pun

Faculty of Science & Technology  
University of Macau  
Macau, China  
{yb17403, cmpun}@umac.mo

**Abstract**— In this paper, we proposed an object tracking method for video stream based on conventional particle filter. Feature vectors are extracted from coefficient matrices of Discrete Cosine Transform (DCT). The feature, as experiment showed, is very robust to occlusion and rotation and it is not sensitive to scale changes. The proposed method is efficient enough to be used in a real-time application. The experiment is carried out on some common used datasets in literature. The result is satisfied and showed the estimated trace follows the target object very closely.

**Keywords**-Object tracking, particle filter, DCT

## I. INTRODUCTION

Since digital camera developed quickly in recent decades, visual object tracking has drawn significant attention from researchers. The visual object tracking techniques could be integrated into many applications like video surveillance, gesture recognition and even augment reality. The aim of object tracking is to estimate the object location from consecutive frames/image sequence and output the trajectory for further analysis.

The framework of tracking system generally includes three parts: 1) Tracking model; 2) Object features to track; 3) Ancillary machine learning approaches, e.g. template updating, for handling object variants. Some approaches might be able to track multiple objects by using object detection methods as well. The simplest tracking models is KL tracker proposed by Lucas and Kanade (K-L tracker) [1]. The K-L tracker makes use of both spatial intensity gradient of the image and Newton-Raphson iteration technique to find the best match. The more frequently used are Kalman filter and particle filter which based on Bayesian inference and sequential Monte Carlo methods[2]. The later proposed particle filters generally outperforms Kalman filters in efficiency and many variants of particle filters emerged in recent years[3]. The a generic particle filter consist of two stages: prediction and update. In the prediction stage, the state posterior probability density function (pdf) is predicted from a measurement from one frame to the next. After that, the weights associated to particles are updated with the latest measurements. Some tracking models are independent to object features. The object features are vectors extracted from the target objects/candidate regions which could represents the characteristics of them. There are many object features in

literature, e.g., Gabor [4], local binary pattern (LBP)[5], Haar-like feature [6, 7], color histograms. The underlying principle of Gabor, Binary Pattern and haar-like features is to project object target or candidate patches on to a set of basis. The coefficient vectors are used in matching stage as feature vectors. In most cases, the color information is considered as statistical information and rendered as histogram. And distances/similarities for these statistical information are generally computed by the Bhattacharyya Coefficient [8]. During long term tracking, target object may change its appearance gradually, making it different from the original object. The occlusions, change of scale or rotation may lead to a mismatch or a drift. To enhance tracking performance and prevent target drift, some methods also incorporate training approaches, like On-line Random forest [9] and On-line boosting [10].

In this paper, we proposed a robust feature extracted from discrete cosine transform (DCT) working along with particle filter. Through extensive experiments, we proved that the feature is very robust to occlusion and rotation. The outline of this paper is as follows. In the first section, an introduction is given and literature is reviewed. Basic framework, tracking model and proposed feature are illustrated in section 2. Details of experiments and its results are presented in section 3. Finally, we made conclusion in last section.

## II. PROPOSED TRACKING METHOD

### A. Particle Filter

Particle filter is Bayesian inference based model which use a Monte Carlo method sample and approximate the target density with a number of particles. The particle filter framework has proven to be effective in tracking problem[3, 11]. The logic is clear. In object tracking, the states of object are considered to be a function of time  $t$  and formulated as stochastic process. Because of the discrete nature of video frames and the states of object in current time only depends on the previous time, the process is simplified to a Markov Chain. The dependence of each pair of consecutive frames reflects on that the state of next time  $t$  does not change dramatically from the previous time  $t-1$ . Specifically, location of object at time  $t$  is within a neighborhood of previous location at  $t-1$ . Particle filter infers current state by using previous state as prior and sampling current likelihood. Each particle is represented by  $\{x_t^i, w_t^i\}_{i=1}^N$ . Particles are

distributed within a neighborhood that the target possibly exists.  $x_t^i$  is the state (in our case is the location) of the particle  $i$  at time  $t$  while  $w_t^i$  is the weight associated to it. The weight represents how important the corresponding particle is and how similar this particle is to the target object. The particle filter model is given by:

$$p(x_t | z_{1:t}) \propto p(z_t | x_t) \sum_{j=1}^N w_{t-1}^j \cdot p(x_t | x_{t-1}^j) \quad (1)$$

The  $w_t^i$  is updated each time by:

$$w_t^i \propto w_{t-1}^i \frac{p(z_t | x_t^i) p(x_t^i | x_{t-1}^i)}{q(x_t^i | x_{t-1}^i, z_t)} \quad (2)$$

The  $w_t^i$  is normalized each time and has a unity sum. The  $q(x_t^i | x_{t-1}^i, z_t)$  is called proposal distribution which draws  $N$  particles  $\{x_t^i\}$ . Proposal distribution draws the particles from the state space.  $p(z_t | x_t^i)$  is the similarity of the  $i$  hypothesis and the target object  $z_t$ . In our case, in order to simplify the tracking system, we employed the bootstrap particle filter and set  $p(x_t^i | x_{t-1}^i) = q(x_t^i | x_{t-1}^i, z_t)$ . The formula (2) becomes:

$$w_t^i \propto w_{t-1}^i \cdot p(z_t | x_t^i) \quad (3)$$

### B. Feature Extraction

For the reason that objects to be tracked are usually dominated by one to three textures and the discrete cosine transform (DCT) coefficient feature is well-accepted as a good feature which accurately describes textures, we employ DCT coefficients as features [12, 13] to find out usefully information that represent objects. Fig.1 shows the framework of our proposed approach which use DCT feature in particle filter framework. Features extracted from DCT coefficients are used in matching the original object patch (target) and the candidate patches (hypothesis) which are the bounding boxes that center particle locations. The similarities between target and candidates are converted into weights for each particle. The location of object at time  $t$  is

estimated by  $x_k = \sum_{i=1}^N w_k^i x_k^i$  with  $\overline{w_k}$  from (3). When locating object in the frame sequence, each frame will be considered an individual image. The DCT is done on intensity map only. For color images, we convert them into grayscale images before transformation.

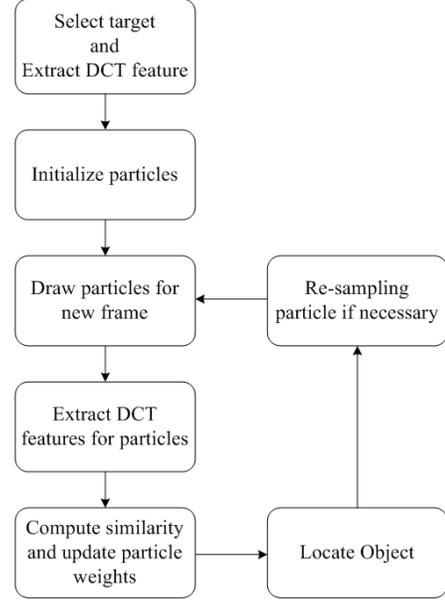


Fig. 1. the framework of our proposed approach.

### III. EXPERIMENTAL RESULTS

Our experiment was carried out in Matlab and evaluated the proposed approach on several video datasets. The objects in tested video sequence are of different degree of non-rigid transforms, including occlusion, rotation and change of scale. Moreover, we did not avoid the changing of lighting condition in the video. The histogram equalization was taken used accordingly as well. Except that “sylv” sequence used 300 particles, the number of particle are set to 100 for all other datasets. Although the experiment is off-line, our method could be implemented in real-time application thanks to the efficiency of our algorithm. The process for each frame is completed within 0.03 second on average, making it capable to handle video stream with 30 frames per second.

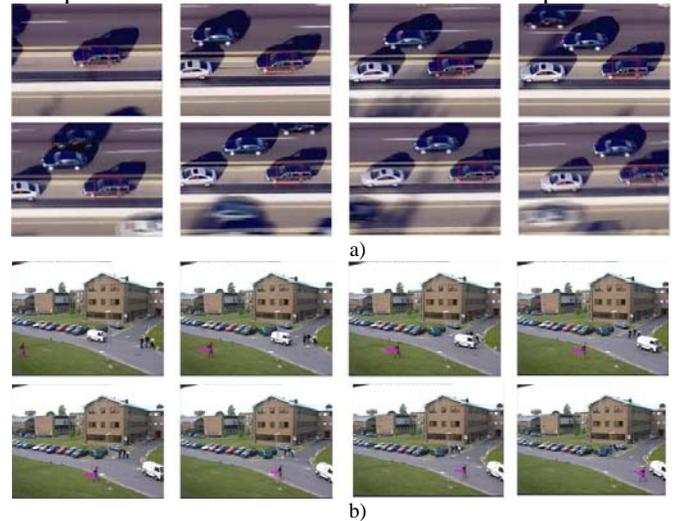


Fig. 2. a) A car runs on the free way. b) A person walks across grasses, with leg movement and direction changes.

The proposed approach was firstly tested on two simple datasets in order to show it is feasible and effective. There are 200 images in the “car\_200” image sequence. The target black car moves fast on a free way. The car changes very little while the surrounding changes dynamically with other cars passing by. Fig.2 a) showed the result for “car\_200” dataset. The red box is manually picked up at the first image. The object was tracked from the first image and the results shown above are snapshots with interval of 20 images. The red box is always following the correct positions of the car. Fig.3 b) is the experimental result for a video clip for PETS2001 dataset. A man walks across the grass and finally enters the sidewalk. The leg movement and change of walking direction cause some object variance. The initial position and the cropping box were obtained from ground truth data. From left to right, snapshots were taken every 31 frames. The pink line is the tracking trace (previous 20 estimated locations) of the target.

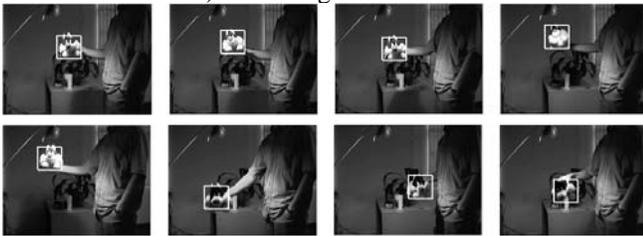


Fig. 3. A doll moves with pose, scale and illumination changes. The numbers of frame showed above are 24, 48, 158, 273, 383, 465, 515 and 606.

#### IV. CONCLUSION

By exploiting representative features from DCT coefficients, we proposed a novel tracking method for video stream based on traditional particle filter. The proposed tracking system is very efficient and is very robust to occlusion and rotation and it is not sensitive to scale changes. Moreover, it is fast enough to be implemented as a real-time tracker. The experimental result is satisfied and showed the estimated trace follows the target object very closely.

#### REFERENCES

- [1] 1. Lucas, B.D., Kanade, T.: An iterative image registration technique with an application to stereo vision. Proceedings of the 7th international joint conference on Artificial intelligence - Volume 2. Morgan Kaufmann Publishers Inc., Vancouver, BC, Canada (1981)
- [2] 2. Arulampalam, M.S., Maskell, S., Gordon, N., Clapp, T.: A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *Signal Processing, IEEE Transactions on* **50** (2002) 174-188
- [3] 3. Hwang Ryol, R., Huber, M.: A particle filter approach for multi-target tracking. *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on* (2007) 2753-2760
- [4] 4. Feng, T., Hai, T.: Non-orthogonal Binary Expansion of Gabor Filters with Applications in Object Tracking. *Motion and Video Computing, 2007. WMVC '07. IEEE Workshop on* (2007) 24-24
- [5] 5. Ojala, T., Pietikainen, M., Maenpaa, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **24** (2002) 971-987
- [6] 6. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, Vol. 1 (2001) I-511-I-518 vol.511
- [7] 7. Lienhart, R., Maydt, J.: An extended set of Haar-like features for rapid object detection. *Image Processing, 2002. Proceedings. 2002 International Conference on*, Vol. 1 (2002) I-900-I-903 vol.901
- [8] 8. Comaniciu, D., Ramesh, V., Meer, P.: Real-time tracking of non-rigid objects using mean shift. *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*, Vol. 2 (2000) 142-149 vol.142
- [9] 9. Saffari, A., Leistner, C., Santner, J., Godec, M., Bischof, H.: On-line Random Forests. *Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on* (2009) 1393-1400
- [10] 10. Leistner, C., Saffari, A., Roth, P.M., Bischof, H.: On robustness of on-line boosting - a competitive study. *Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on* (2009) 1362-1369
- [11] 11. Xue, M., Haibin, L.: Robust Visual Tracking and Vehicle Classification via Sparse Representation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **33** (2011) 2259-2272
- [12] 12. Grissom, E.A., Rajan, P.K.: An algorithm for a fast two-dimensional discrete cosine transform. *System Theory, 1990., Twenty-Second Southeastern Symposium on* (1990) 646-651
- [13] 13. Wen-Hsiung, C., Smith, C., Fralick, S.: A Fast Computational Algorithm for the Discrete Cosine Transform. *Communications, IEEE Transactions on* **25** (1977) 1004-1009
- [14] 14. Babenko, B., Ming-Hsuan, Y., Belongie, S.: Visual tracking with online Multiple Instance Learning. *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on* (2009) 983-990