ADAPTIVE RESOURCE ALLOCATION IN PARTICLE FILTERING FOR ARTICULATED OBJECT TRACKING

Pan Pan and Dan Schonfeld

ECE Department, University of Illinois at Chicago
Chicago, IL, USA, 60607 Email:{ppan3, dans}@uic.edu

ABSTRACT

This paper presents a novel particle allocation approach to particle filtering for articulated object tracking which minimizes the total tracking distortion given a fixed number of particles over a video sequence. Under the framework of decentralized articulated object tracking, we propose the dynamic proposal variance and optimal particle number allocation algorithm for articulated object tracking to allocate particles among *different parts of the articulated object* as well as *different frames*. Experimental results show the superior performance of our proposed algorithm to traditional particle allocation methods, i.e. a fixed number of particles for each object part in each frame. To the best of our knowledge, our approach is the first to provide an optimal allocation of a fixed number of particles among different object parts and different frames.

***Index Terms*—** Tracking, resource management, particle allocation, articulated object tracking, rate distortion theory

1. INTRODUCTION

Over the past decade, particle filters have gained enormous popularity in video tracking [1]. In the particle filtering framework, a proposal density is used to generate samples. The samples are used to evaluate the importance weights, which are normalized and subsequently used to the estimate the posterior density function. There exists a tradeoff between tracking quality and tracking computational resources. On one hand, the sample approximation of the posterior density can be made arbitrarily accurate if the number of particles is sufficiently large. On the other hand, the number of particles used is an essential index of the complexity of the implementation of particle filters. In articulated object tracking (AOT), the tradeoff becomes more severe than that in single object tracking because of the high degrees of freedom of the object.

Recent technological trends have required the deployment of particle filters for video tracking applications in mobile devices, e.g. handheld video phones. The limited power and scarce computational resources available in embedded computer systems have imposed tremendous constraints on the number of particles used for tracking.

Past research of AOT has been emphasized on the interaction between different object parts. The common approach to address this problem is a joint state space representation. However, the number of particles it demands grows exponentially in terms of the degrees of freedom of the object. Qu et al. [2] has demonstrated that a decentralized Bayesian framework can be used to maintain a linear increase in the number of particles as the number of object parts increase.

Even though under the decentralized framework, we should use the limited number of particles wisely in order to achieve the best tracking quality. Traditionally, the proposal variance and the number of particles per object part per frame are fixed during the entire tracking process. These parameters are set based on trial-and-error experiments prior to tracking. However, this approach does not consider the different characteristics of each object part in each frame. For example, within one frame, some parts of the articulated object move fast, other parts move slowly or stay still. For the same part of the articulated object, it may move fast in some frames and slow down in the next few frames. When computing and power resources are limited, we should use these information to utilize the available resources wisely and attain the best tracking quality possible. MacCormick and Isard [3] presented survival diagnostic and survival rate as quantities to assess the efficacy of particles filters. However, those concepts cannot tell how to allocate particles between partitions in partitioned sampling. In our previous work, we proposed approaches to do particle allocation for single object tracking [4], as well as for multiple object tracking [5]. In this paper, under the framework of decentralized articulated object tracking [2], we exploit the characteristic behavior of the object parts to dynamically vary the proposal variance and allocate the optimal number of particles for each object part as well as for each video frame.

The rest of the paper is organized as follows, in Section 2 we propose a novel criteria to measure the efficiency of particle filtering and derive the optimal particle allocation equation for decentralized articulated object tracking. We then introduce dynamic proposal variance and optimal particle allocation algorithm in Section 3. The experimental results with comparison to other methods are given in Section 4, followed by the conclusion given in Section 5.

2. OPTIMAL RATE-DISTORTION

2.1. Tracking Distortion

In particle filtering, let the tracking error ϵ^i of the i^{th} particle be defined as the difference between the true state S and the sampled state X^i, i.e. $\epsilon^i = S - X^i$. Then we define the total tracking error Y as the difference between true state S and estimated state $\hat{X}$, i.e. $Y = S - \hat{X}$, where S, $\hat{X}$ and Y are vectors with dimension N in video tracking. If the number of particles n is sufficiently large, it can be shown that Y is a zero-mean random vector (i.e. $E(Y) = \mathbf{0}$). This complies with the fact that the estimation of the mean given by particle filter is asymptotically unbiased. We therefore define tracking distortion D as the variance of the total tracking error Y, i.e.

$$D = var(Y) = E[Y^T Y] = \sum_{l=1}^{N} var(Y^l) = \frac{\sigma^2 \zeta}{n}, \qquad (1)$$

where $\sigma^2 = \sum_{l=1}^{N} \sigma_l^2$ and σ_l^2 is the variance of the l^{th} component of ϵ^i; $\zeta = E(\frac{2\epsilon_{max} \int_{-\epsilon_{max}}^{\epsilon_{max}} w^2(\epsilon)\beta(\epsilon)d\epsilon}{(\int_{-\epsilon_{max}}^{\epsilon_{max}} w(\epsilon)\beta(\epsilon)d\epsilon)^2})$ is a constant and $w(\cdot)$ is the un-normalized weight function; β is the interval density of ϵ^i; ϵ_{max} is the error bound of each particle.

Equation (1) corresponds to the result of the convergence of the variance of the particle filter estimator in [6]. From (1), we observe that for fixed $\sigma^2\zeta$, the tracking distortion D decreases as the number of particles increases. As the number of particles n tends to infinity, the tracking distortion D approaches zero. This observation is consistent with the theory of Bayesian tracking.

2.2. Problem Formulation

We now use rate distortion theory to derive the optimal particle allocation equation for articulated object tracking. Under the framework of decentralized articulated object tracking (DAOT) [2], each part of the articulated object is tracked by a decentralized tracker. We define the total tracking distortion *per frame* as the average of the distortion of all the object parts in one frame. The total tracking distortion *over a video sequence* is the average of the distortion in all frames of this video sequence.

Since we want to attain the best tracking quality possible by minimizing the tracking distortion, we consider a constraint on the average number of particles n used over J frames and K object parts. Then our problem can be interpreted as: to determine the optimal number of particles $n_{j,k}$ for the k^{th} part in the j^{th} frame by allocating the total of nJK particles among J frames and K parts so that the total distortion D_T is minimized, i.e.

$$\min\ D_T = \frac{1}{JK} \sum_{j=1}^{J} \sum_{k=1}^{K} \frac{\sigma_{j,k}^2 \zeta_{j,k}}{n_{j,k}} \ \text{ s.t. } \sum_{j=1}^{J} \sum_{k=1}^{K} n_{j,k} = nJK.$$

We solve this constrained optimization problem by forming the Lagrangian P given by

$$P = \sum_{j=1}^{J} \sum_{k=1}^{K} \frac{\sigma_{j,k}^2 \zeta_{j,k}}{n_{j,k}} + \lambda(\sum_{j=1}^{J} \sum_{k=1}^{K} n_{j,k} - nJK).$$

By setting $\partial P / \partial n_{j,k} = 0$, we obtain

$$n_{j,k} = n \frac{\sqrt{\sigma_{j,k}^2 \zeta_{j,k}} JK}{\sum_{j'=1}^{J} \sum_{k'=1}^{K} \sqrt{\sigma_{j',k'}^2 \zeta_{j',k'}}}. \qquad (2)$$

The parameter $\zeta_{j,k}$ is difficult to compute. However, under the assumptions that ϵ_{max}, $\beta(\cdot)$ and $w(\cdot)$ of each object part of adjacent frames are approximately the same, $\zeta_{j,k}$ will be independent of frame j and part k. Therefore, we observe that (2) is given by

$$n_{j,k} = n \frac{\sqrt{\sigma_{j,k}^2} JK}{\sum_{j'=1}^{J} \sum_{k'=1}^{K} \sqrt{\sigma_{j',k'}^2}}. \qquad (3)$$

Although (3) is valid under certain assumptions, it provides a good approximation which captures the relationship between $n_{j,k}$ and $\sigma_{j,k}^2$. The assumptions imposed allow us to use this equation in practical algorithms. From (3), we observe given the average number of particles n used over J frames among K object parts, the number of particles $n_{j,k}$ allocated to the k^{th} part in the j^{th} frame is determined by the error variance $\sigma_{j,k}^2$. An object part with a large error variance should be allocated more particles; whereas a part with a smaller error variance should be assigned fewer particles. We finally obtain the optimal distortion of the k^{th} part in the j^{th} frame as $D_{j,k} = \frac{\sqrt{\sigma_{j,k}^2 \zeta_{j,k}}}{nJK} \sum_{j'=1}^{J} \sum_{k'=1}^{K} \sqrt{\sigma_{j',k'}^2 \zeta_{j',k'}}$.

3. OPTIMAL PARTICLE ALLOCATION FOR ARTICULATED OBJECT TRACKING

3.1. Error Variance and Proposal Variance

In particle filtering, the particles are generally sampled using a sampling scheme given by $X_j^i = f(X_{j-1}^i) + v_j$, where $f(X_{j-1}^i)$ can be any estimation of the mean of the new sample and v_j is given by a Gaussian distribution $v_j \sim \mathcal{N}(0, \mathbf{R}_G)$, where $\mathbf{R}_G = diag(\varphi_1^2, \varphi_2^2, \ldots, \varphi_N^2)$. The variance of the l^{th} component of v_j is called proposal variance φ_l^2. Due to the effect of resampling and the choice of the proposal density, we can view the variance of $f(X_{j-1}^i)$ is approximately zero. Therefore, we observe that for each component of the tracker for each object part in the j^{th} frame

$$\sigma_l^2 = var(\epsilon_j^{i,l}) = var(S_j^l - X_j^{i,l}) = var(v_j^l) = \varphi_l^2,$$

for $l = 1, 2, \ldots, N$. Therefore, the tracking error variance σ_l^2 is equal to the proposal variance φ_l^2 used in particle filtering.

3.2. Dynamic Proposal Variance

In traditional implementation of DAOT, the proposal variance $\varphi_{j,k}^2$ is fixed for all object parts in all frames, i.e. $\varphi_{j,k}^2 = \varphi^2$. From (3), we observe that in this case the optimal number of particles $n_{j,k}$ is uniform for all object parts in all frames, i.e. $n_{j,k} = n$. The proposal variance is selected manually prior to tracking for different video sequences. However, the method of using a fixed proposal variance for all object parts and all frames fails to exploit the different characteristics of each object part in each frame to improve the sampling scheme. For example, we may wish to sample fast moving object parts using a proposal density function with a larger variance. We therefore introduce the dynamic proposal variance given by

$$v_{j,k} \sim \mathcal{N}(0, \mathbf{R}_{\Delta \widehat{X}_{j,k}}). \tag{4}$$

This scheme implies that the variance of the proposal density $\varphi_{j,k}^2$ is changing with the estimated object parts motion $\Delta \widehat{X}_{j,k}$. The optimal number of particles $n_{j,k}$ will be allocated to each object part in each frame according to the proposal variance as in (3).

We will rely on the motion vectors to determine the variance of the position components in the proposal density function. Given the motion vector $(\Delta x, \Delta y)$, the variance should ensure that the position of the object part in the next frame lies in the search region of the current frame. Therefore, we obtain

$$\varphi_x^2 = c\sqrt{2}\Delta x, \varphi_y^2 = c\sqrt{2}\Delta y, \tag{5}$$

where c is a constant and (x, y) is the center of the object part. In practice, we let $c = 1 \sim 2$ for the sampling scheme $X_{j,k}^i = X_{j-1,k}^i + v_{j,k}$, and we choose $c = 0.1 \sim 0.2$ for the sampling scheme $X_{j,k}^i = X_{j-1,k}^i + \Delta \widehat{X}_{j,k} + v_{j,k}$. We could also adjust the variance of other dimensions of the proposal density, e.g. zooming and rotation, based on the principles presented above.

3.3. Optimal Particle Allocation (OPA) Algorithm

Let us assume that we process J frames at a time. When the time elapsed during J frames is only a small fraction of a second, we can consider the proposed approach for real-time tracking systems. The optimal particle allocation (OPA) algorithm for articulated object tracking is illustrated as:

1. Use motion detection scheme (e.g. block matching) or learned dynamics to estimate $\Delta \widehat{X}_{j,k}$ for each object part in each frame.

2. Choose the proposal variance $\varphi_{j,k}^2$ dynamically according to motion $\Delta \widehat{X}_{j,k}$.

3. Use (3) to determine the optimal particle number allocated to each object part in each frame.

4. Do decentralized articulated object tracking based on the number of particles allocated.

5. Repeat steps 1-4 for each group of J frames throughout the entire video sequence.

4. EXPERIMENTAL RESULTS

In articulated object tracking, we are more interested in particle allocation among different object parts in each frame, i.e. $J = 1$. To demonstrate the improved performance of the proposed OPA algorithm, all of the comparative experiments are performed under the DAOT framework with the sampling scheme $X_{j,k}^i = X_{j-1,k}^i + v_{j,k}$ and $J = 1$. For the implementation details of the DAOT algorithm, we refer the reader to [2]. The pose relation and motion are learned from the training data. Each object part is modelled by a four dimensional rectangular with different color for labelling. We use color characteristic only as cue for computing the local particle likelihood. We use inter-part interactive model to deal with interaction between adjacent parts.

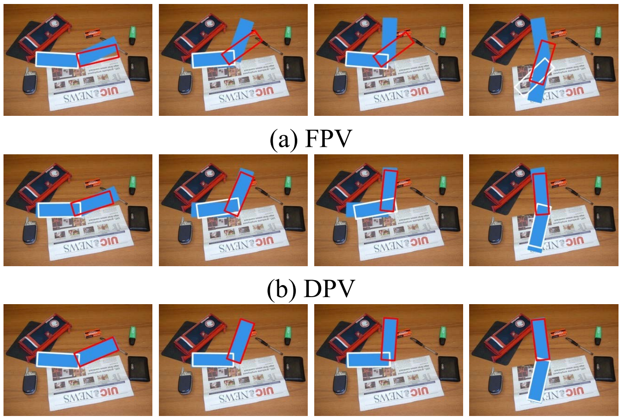
We compared our approach with two other variance choosing and particle allocation algorithms on both synthetic and real video sequences: (a) Fixed Proposal Variance, Fixed Particle Allocation (FPV). This is the traditional implementation of DAOT. In practice, the variance is set before tracking to an arbitrary value. In the following experiments, we set the variance to the average value of the dynamic variances obtained by our method. (b) Dynamic Proposal Variance, Fixed Particle Allocation (DPV). This is a modification of FPV where the variance of the proposal density function is dynamically adjusted. (c) Dynamic Proposal Variance, Optimal Particle Allocation (OPA). This is the proposed algorithm.

The synthetic sequence `Cardboard` has two identical cardboard jointly moving in a challenging clutter environment with a resolution of 320×240. At each time, only one cardboard ball moves fast, the other one moves slowly. The average number of particles per object part per frame n is 30. Tracking results of the different algorithms are shown in Fig. 1. Our proposed OPA algorithm outperforms than other two algorithms while requiring about the same CPU time (see Table 1).

The `Boy` video clip contains a boy moving each one of his arms at one time. The resolution is 320×240 and the frame rate is 20 frames per second. The average number of particles n is 10 per object part per frame. Fig. 2 illustrates the tracking results of the different algorithms, with the actual number of particles each object part used shown in Fig. 2 (c). Our OPA algorithm produces better tracking results compared to the other algorithms.

We have implemented all of the algorithms independently in VC++ without code optimization on a 2.8 GHz Pentium IV PC. Since the total number of particles used per frame is the same, we can check from Table 1, the computation time per frame of the three algorithms is the same. The data has been averaged over 5 iterations on the `Cardboard` sequence.

FPV cannot adjust the proposal variance dynamically according to the motion of the object. In practice, its value must be set to a predetermined value before tracking. Low variance will cause tracking to be completely lost, whereas a large variance will cause jittering and shaking. A suitable fixed proposal variance for all parts of the video and for all kinds of the videos cannot be found. DPV can adjust the proposal variance based on motion, and therefore it can be used in all tracking systems. However, when the computational resources are limited, i.e. the total number of particles available is fixed, we should also allocate particles according to the dynamic proposal variance in order to maximize the tracking quality as in OPA. Hence, OPA can give the best performance among these three.

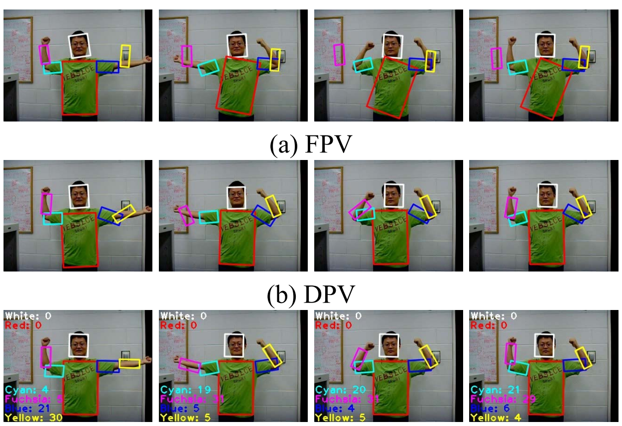


(a) FPV

(b) DPV

(c) OPA

Fig. 1. Tracking results of the synthetic video `Cardboard`: (a) FPV, (b) DPV and (c) OPA. (Each row depicts frames 25, 72, 89, 177.)



(a) FPV

(b) DPV

(c) OPA

Fig. 2. Tracking results of the real video `Boy`: (a) FPV, (b) DPV and (c) OPA. (Each row depicts frames 93, 290, 466, 508.)

Table 1. Computation Time Per Frame On `Cardboard` Sequence

FPV	DPV	OPA
16 millisecond	16 millisecond	16 millisecond

5. CONCLUSION

In this paper, we presented a new approach for articulated object tracking which minimizes the total tracking distortion by simultaneously adjusting the proposal variance and the number of particles for each object part in each frame based on the motion activity of the tracked object. We derived a theoretical framework based on rate distortion theory to determine the optimal number of particles allocated for each part in each frame. The motivation of our approach is to maximize the performance of particle filters in applications that have limited computational and power resources. Our proposed optimal particle allocation (OPA) algorithm has the following advantages: (1) it can minimize the total tracking distortion while using the same number of particles; (2) given a fixed power, it can achieve the best tracking quality; (3) for the same tracking quality, it uses the least CPU time and power.

6. REFERENCES

[1] P. Perez, J. Vermaak and A. Blake, "Data Fusion for Visual Tracking with Particles," *Proceedings of the IEEE*, vol. 92, no. 3, 2004, pp. 495-513.

[2] W. Qu and D. Schonfeld, "Real-time decentralized articulated motion analysis and object tracking from videos," *IEEE Transactions on Image Processing*, vol. 16, no. 8, 2007, pp. 2129-2138.

[3] J. MacCormick and M. Isard, "Partitioned sampling, articulated objects, and interface-quality hand tracking," *European Conference on Computer Vision*, 2000.

[4] P. Pan and D. Schonfeld, "Adaptive particle-distortion tradeoff control in particle filtering for video tracking," *IEEE International Conference on Image Processing*, 2006, pp. 565-568.

[5] P. Pan and D. Schonfeld, "Optimal particle allocation in particle filtering for multiple object tracking," *IEEE International Conference on Image Processing*, 2007, pp. I-61-I-64.

[6] D. Crisan and A. Doucet, "A survey of convergence results on particle filtering methods for practitioners," *IEEE Transactions on Signal Processing*, vol. 50, no. 3, 2002, pp. 736-746.