The Thirteenth International Symposium on Artificial Life and Robotics 2008(AROB 13th '08),
B-Con Plaza, Beppu, Oita, Japan, January 31-February 2, 2008

Adaptive Particle Allocation for Multifocal Visual Attention based on Particle Filtering

Naomi Yano,[1] Tomohiro Shibata[1], Shin Ishii[2]

[1]Graduate School of Information Science, Nara Institute of Science and Technology (NAIST)
8916-5 Takayama, Ikoma, Nara, 630-0192
[2]Graduate School of Informatics, Kyoto University
Gokasho, Uji, Kyoto, 611-0011

Abstract

Confronting flood of visual inputs, examining all possible interpretations based on the given visual data is impossible in principle. Despite of these computational problems, humans robustly perform visual processing accurately. One of the most important keys in the human visual processing would thus be attention control.

In this article, we first indicate that Particle Filter (PF) is a major candidate for the model of multifocal visual attention. PF is a method which approximates intractable integrations in incremental Bayesian computation by means of stochastic sampling. One of the major drawbacks of PFs is a trade-off between computational costs and tracking performance; a large number of particles are required for accurate and robust estimation of state variables, which is time-consuming. This study proposes a computational model for multifocal visual attention which deals with the cost-performance trade-off with the restricted computing resource (the number of particles). Simulation experiments of tracking two targets with only tens of particles demonstrate the feasibility of the model.

1 Introduction

There exists inevitable ill-posedness in the visual information provided by the environment. Confronting flood of visual inputs, examining all possible interpretations based on the given visual data is impossible in principle. Despite of these computational problems, humans robustly perform visual processing accurately. One of the most important keys in the human visual processing would be attention control. From the computational viewpoint, attention control is a mechanism to enable the human to perform real-time acquisition of a meaningful solution (interpretation) by combining some *a priori* knowledge, prediction, or hypothesis of the target, and observed data, and by actively ignoring irrelevant data. Classical theories of attention assumed a single focus within the visual field, but previous psychophysical studies have indeed demonstrated the human's ability to simultaneously track four or more targets in the visual field [1]. The mechanisms by which attention maintains focus and assigns its cognitive resource on several targets are, however, not yet established.

In this study, we propose a computational (engineering) model of multifocal visual attention based on Particle Filters (PFs). The rest of this article is organized as follows. In the next section, we first explain why PFs can be a major candidate for the model of multifocal visual attention. Section 3 introduces our model. Section 4 describes simulation results. Some discussion including possible future works are done in section 5.

2 Modeling Human Visual Attention by Particle Filters

There have been many behavioral and computational studies reporting that the brain would compute Bayesian statistics [2]. Performing Bayesian estimation of posterior distributions is, however, intractable in general. Particle filtering is an approach to performing Bayesian estimation of intractable posterior distributions from time-series signals with non-Gaussian noise [3]. It has been studied in various research areas, including real-time visual processing, which deals with general cases in which images are contaminated by non-Gaussian noise due not only to signal noise but also to the existence of obstacles and/or distracters [4]. Furthermore, PFs are easy to be implemented and parallelizable. These attractive features seem to be shared with neural computations in the brain.

We regard PF as a model of attention control because (1) it employs sequential importance sampling (SIS), and (2) it shares the same critical computational problem with human visual processing. SIS is nothing but a recursive version of importance sampling which is a generally used

The Thirteenth International Symposium on Artificial Life and Robotics 2008(AROB 13th '08),
B-Con Plaza, Beppu, Oita, Japan, January 31-February 2, 2008

technique in Monte-Carlo integration. The importance sampling distribution is used to determine the weight of each partcile, and works like attention, e.g., if the distribution is spatially localized, it attends the corresponding region, whereas ignores other regions.

There are two major drawbacks of PFs, a naive resolution of each of them requires a large number of particles, which is time-consuming. The first drawback occurs when the importance sampling distribution does not match the true distribution. In other words, errors in the prior knowledge, or the model, make many particles have insignificant weights, leading to misestimation. The second drawback occurs when the dimension of the state variables is high, i.e., curse of dimensionality. In this study, we focus on dealing with the first drawback, while the second drawback is just solved by assuming that the state variables are mutually independent.

Here, we consider tracking a single target. At time t, suppose the target state $\boldsymbol{x}_t$ is generated by Eq. (1) and observed as $\boldsymbol{z}_t$ according to Eq. (2), where A,B,C and D are state transition matrix, standard deviation vector to control the strength of system noise, observation vector, standard deviation of the observation noise, respectively, and the noises u_t and v_t follow one-dimensional normal Gaussian distributions,

$$\boldsymbol{x}_t = A\boldsymbol{x}_{t-1} + Bu_t, \quad (1)$$
$$z_t = C\boldsymbol{x}_t + Dv_t. \quad (2)$$

In the simulation, we set the parameters of the true model as

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}, B = \begin{pmatrix} b \\ 0 \end{pmatrix},$$
$$C = \begin{pmatrix} 1 & 0 \end{pmatrix}, D = 0.1,$$

and $b = 0.1$. Fixing A means that the target velocity is constant. Instead of using these true parameters, an A' different from A was used for particle filtering of the state variable $\boldsymbol{x}_t$, as to represent the prior knowledge of the target motion:

$$A' = \begin{pmatrix} \cos(\omega T) & \frac{1}{\sin(\ T)} \\ \sin(\omega T) & \cos(\omega T) \end{pmatrix}, \quad (3)$$

indicating that the target motion is assumed to be sinusoidal with the angular velocity of $\omega = 2\pi$ [rad/s] and the sampling period of $T = 0.05$ [s].

We examined tracking accuracy when the number of particles N was varied, from 10 to 100 by 10, and 30 trials were performed for each N. Tracking errors were obtained by absolute difference between the true value and the estimated value, Fig. 1. In this figure, median of each tracking error is shown. The blue line is for when PF knew the true model. The green line is for when the A' was used in PF.

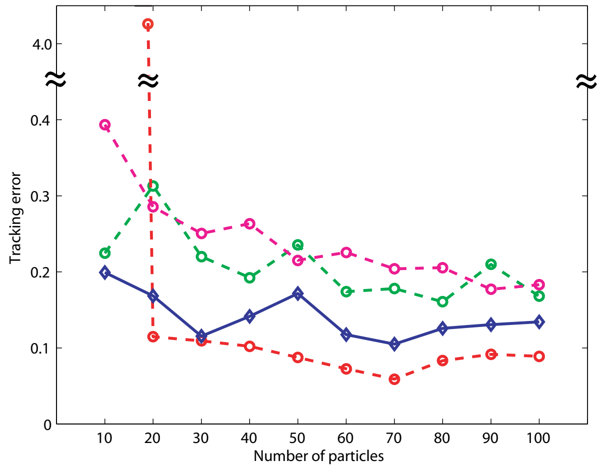

Figure 1: Effects of particle numbers for single tracking

The red and purple lines are for when the b was set to 0.9 and the D' was set to 0.9, respectively.

This figure demonstrates that there is a trade-off between computational cost and tracking performance, which also exists in the human visual processing, giving strong motivation for modeling human visual attention by particle filtering.

3 Adaptive Particle Allocation

As a solution to the trade-off problem described in the last section, we propose an adaptive particle allocation (APA) method. The idea of APA is as follows. If there are multiple targets which have comparable tracking difficulties but have no priority for tracking, then the restricted computational resource (the number of particles in our study) should be allocated to them equally. In contrast, if the targets have different tracking difficulties, the number of particles can be differently allocated to each target. In this study, we assume that each target has unreliability index of tracking, and APA allocates the number of particles to the targets according to the index. The tracking difficulty is represented by the model error described in the previous section. The unreliability U_i of the target $i (i = 1, ..., M)$ is then defined by

$$U_i = \frac{\sigma_i^2}{\sum_{k=1}^{M} \sigma_k^2}, \quad (4)$$

where σ_i^2 is the estimated variance of the target i. Since the unreliability is expected to be low during successful tracking, and vice versa, we use the following simple rule for allocating the number of partiles N_i for each target i:

$$r_i = U_i + \ , \quad (5)$$
$$N_i = r_i N. \quad (6)$$

That is, for target i, N_i is allocated in proportional to its tracking unreliability (cf. Fig. 2).

Because $\sum_i^M N_i = N$, there is a constraint for r_i that $\sum_i^M r_i = 1$. Then, and have also a constraint as follows.

$$\sum_{i=1}^{M} r_i = \sum_{i=1}^{M} \quad U_i + \quad = \quad + M \quad = 1. \qquad (7)$$

Here, we use the fact that $\sum_i^M U_i = 1$.

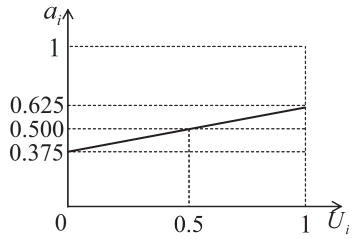

Figure 2: Example allocation function

Fig. 3 summerizes the proposed PF algorithm with APA. Fig. 2 depicts an example allocation function in a case that $= 0.25$ and $= 0.375$. In this case, the particle allocation rate r_i can vary only within a narrow region ($0.375 \quad r_i \quad 0.625$). In the next section, we indeed use this range of the particle allocation rate due to the following reason. Because we focus on the case where the number of particles is small, allowing the rate close to one or zero is impractical. Namely, if we allow the wider range, the allocated number of a target can be close to zero, e.g., $N = 10$ and r_i=0.1 leads to $N_i = 1$.

4 Simulation Experiments

4.1 Experimental Setup

Experiments of multi-target tracking were carried out to investigate the feasbility of APA. We investigated the case where there are two targets independently moving in a one-dimensional space, and there is no interefere between their motions. Their appearance is assumed to be completely distinguishable, i.e., there is no interfere between their observations. Their motions are governed by linear-Gaussian state-space representation shown as Eqs. (1)(2). The tracking performances were compared in the case APA was used to the case APA was not used under the assumption that the particle filter used wrong models of state-transition and observation for target 2.

In the experiments, the following parameter settings were used for the two targets $i(i = 1, 2)$:

$$A_1 = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}, A_2 = \begin{pmatrix} \cos(\omega T) & \frac{1}{\sin(\ T)} \\ \sin(\omega T) & \cos(\omega T) \end{pmatrix},$$

$$B_1 = \begin{pmatrix} b_1 \\ 0 \end{pmatrix}, B_2 = \begin{pmatrix} b_2 \\ 0 \end{pmatrix}, D_1 = D_2 = 0.1,$$

and $b_1 = b_2 = 0.1$. The angular velocity was $\omega = 2\pi$ [rad/s] and the sampling time was $T = 0.05$ [s]. $= 0.25$ and $= 0.375$ were used throught the experiments.

Perform the followings for each target $i(i = 1, ..., M)$ and for time step $t(t = 1, \quad ,)$

- predict by sampling from $p(\boldsymbol{x}_{i\ t}|\boldsymbol{x}_{i\ t-1} = \boldsymbol{s}_{i\ t}^{(n)})$
- observe and weight the state in terms of $\boldsymbol{z}_t$.
$$w_{i\ t}^{(n)} = p(\boldsymbol{z}_t|\boldsymbol{x}_{i\ t} = \boldsymbol{s}_{i\ t}^{(n)}) \qquad (8)$$
- estimate mean and variance of the weighted state
$$x_{i\ t} = \frac{\sum_{n=1}^{N}(w_{i\ t}^{(n)} x_{i\ t}^{(n)})}{\sum_{k=1}^{N} w_t^{(k)}} \qquad (9)$$
$$\sigma_i^2 = \frac{(w_{i\ t}^{(n)}(x_{i\ t}^{(n)} \quad x_{i\ t}^{(n)})^2)}{\sum_{n=1}^{N} w_{i\ t}^{(n)}} \qquad (10)$$
- APA: calculate unreliability and determine N_i for each target i ($i = 1, ..., M$)
$$U_i = \frac{\sigma_i^2}{\sum_{k=1}^{M} \sigma_k^2} \qquad (11)$$
$$r_i = \quad U_i + \quad , \qquad (12)$$
$$N_i = r_i N \qquad (13)$$
- resample N_i particles $\{\boldsymbol{s}_{i\ t+1}^{(n)}\}$ from the sample-set $\{\boldsymbol{s}_{i\ t}^{(n)}, \boldsymbol{w}_{i\ t}^{(n)}\}$ for each target i

Figure 3: Particle Filtering with APA

4.2 Results

Fig. 4 presents example tracking results with $N = 70$ in a case that the state transition model of the two targets was assumed to be constant but was wrong for the target 2. The upper panel shows successful tracking of both the targets. The observation and estimated values of the target1 / 2 were illustrated respectively by red / green solid line and blue / black dash line. The middle and lower panels show that the high variance sometimes occured for tracking the target 2, but it was eased by APA.

Fig. 5 shows tracking results for various numbers of particles, in a particular case that the state transition model of the two targets was assumed to be constant but was wrong for the target 2. The left panel (a) shows the tracking results of target 1, while the right panel (b) target 2. As shown in this figure, tracking accuracy was worse without APA, which is more evident when the number of particles is smaller.

Table 1 compares median tracking errors of target 2 with APA and to those without APA for six different total numbers of particles and in three different cases. More specifically, each value is a median tracking error without APA divided by the one with APA. Therefore, when the value is larger than one, the median tracking error with APA is smaller than that without APA. In Case 1, the state transition model of the two targets was assumed to be constant but was wrong for the target 2. In Case 2, the system noise b_2 was set to 0.9 which was quite different from the true system noise $b = 0.1$. In Case 3, the observation noise D_2 was set to 0.9 which was also quite different from the actual observation noise $D = 0.1$. Each tracking error was a median error of 30 trials. As shown in this table, tracking accuracy was worse without APA, which is more evident when the number of particles is smaller.

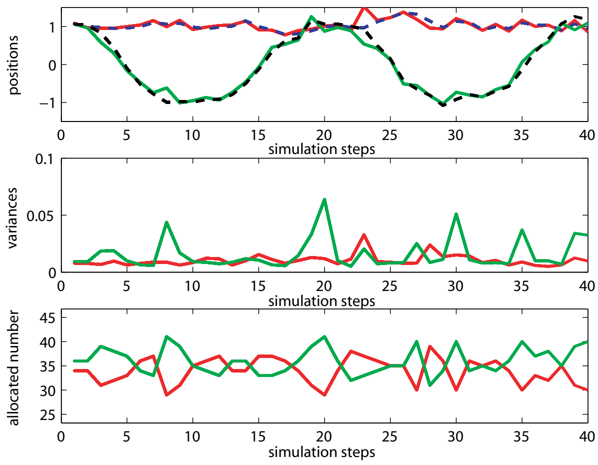

Figure 4: Example tracking results

Table 1: Tracking errors with and without APA for various total numbers of particles

	20	30	40	50	60	70
Case 1	1.46	1.20	1.01	0.935	4.45	0.975
Case 2	1.07	0.947	1.00	1.03	1.03	1.04
Case 3	1.23	1.02	1.12	0.971	1.06	1.03

5 Conclusion

In this article, we proposed a computational model for multifocal visual attention based on particle filtering. The key of the model is adaptive particle allocation (APA) in which the number of particles allocated for each target was varied according to the unreliablity for each tracking. Although simulation experiments demonstrated the feasibility of APA, a couple of problems remain. First, the parameters for the allocation function were heuristically determined. Those parameters should be optimized as to best

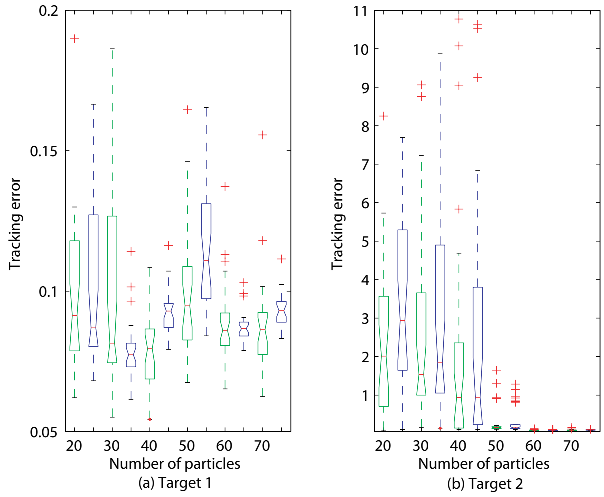

Figure 5: Tracking errors with and without APA. N was varied from 20 to 70. For each number of particles, there are two box plots; the left and green plot shows the result with APA and the right and blue plot without APA.

match the environments, like humans do. In addition, tuning those parameters in an on-line fashion would be important [5]. Investigating the mechanisms of multifocal attention in human vision under the consideration of correspondence to the model proposed in this study is in our future study direction.

References

[1] P. Cavanagh and G.A. Alvarez. Tracking multiple targets with multifocal attention. *TRENDS in Cognitive Science*, 9(7), 2005.

[2] K. Doya, Ishii Shin, A. Pouget, and R.P.N. Rao, editors. *Bayesian Brain: Probabilistic Aproaches to Neural Coding*. MIT Press, 2007.

[3] N.J. Gordon, J.J. Salmond, and A.F.M. Smith. Novel approach to nonlinear non-gaussian bayesian state estimation. In *IEEE Proceedings on Radar Signal Processing*, volume 140, pages 107–113, 1993.

[4] Michael Isard and Andrew Blake. Condensation-conditional density propagation for visual tracking. *International Journal of Computer Vision*, 29(1):5–28, 1998.

[5] T. Shibata, T. Bando, and S. Ishii. Visual tracking achieved by adaptive sampling from hierarchical and parallel predictions. In *ICONIP2007*, 2007. to appear.