A Particle Filter for Tracking a Firefly

Yoshihiko Tsukamoto and Yusuke Matsumoto
Wakayama University
{*tsukamoto,matsugen*}*@vrl.sys.wakayama-u.ac.jp*

Toshikazu Wada
Wakayama University
twada@ieee.org

Abstract

The Particle filter estimates a probability distribution of target object's state by sampled hypotheses and their weights. This method is more expressive than existing method such as Kalman filtering, because the object state is represented as a multi-modal distribution. However, the method can't be directly applied to temporally variable appearance object tracking, for example, a firefly, or a flickering neon-sign. For solving this problem, we propose a particle filter for a variable appearance object, which estimates a unique state parameter independent of target's position. Our method decomposes the state space into disjoint parameter spaces, i.e., object position and posture space and appearance parameter space. In the appearance parameter space, the likelihood of each hypothesis is evaluated at the position parameters generated in the other space, and the best explain parameter is determined. Based on this parameter, likelihood in the position and posture space is evaluated. By interacting the parameter estimations in different spaces, we can successfully track blinking firefly in the darkness.

1 Introduction

The Particle filtering (or called the CONDENSATION [1]) or its extensions generate many hypotheses in a state parameter space spanned by target's position, posture, and shape. The hypotheses' weights are calculated by consistency between hypotheses and captured image [4, 6, 7, 8]. It develops characteristics 1) there are relationships among state parameters, 2) it allows ambiguity about many combinations of parameters. These characteristics should appear as an advantage, however, it may become defects sometimes. For example, in a case when a set of state parameters can be decomposed into some classes, many hypotheses are generated on unnecessary state parameter. In the matter of tracking efficiency, this can become a big problem. For solving this problem, importance sampling [2], partitioned sampling [5] and many methods have been proposed [3].

This problem adversely affects not only efficiency but also tracking stability. For example, a problem of blinking firefly tracking has a parameter including appearance independent of position. The appearance parameter can't be represented as a shape or other geometrical model in the dark scene. Under this condition, we have to estimate a likelihood of firefly from the brightness. Therefore, a likelihood estimation will vary whether each hypothesis assumes "a firefly is lighting" or "disappear." Because, most hypotheses assume "disappear" using original particle filter, the tracking may be performed in a dark place.

For solving this problem, we change criteria of likelihood evaluation depending on unique object appearance parameter. This is equivalent to daringly introducing an uni-modal distribution that does not allow ambiguity into the particle filter allowing ambiguity.

When the ambiguity is excluded completely, the particle filter doesn't have any advantages. Therefore, for tracking variable appearance object, state parameters are considered as two types of parameters; ambiguity or unique. For example, in the case of blinking firefly, the unique parameter represents the brightness and it keeps changing, which is independent of position. For estimating the brightness parameter, we estimate probability distribution which is excluded multimodality. This uni-modal distribution estimation can obtain both a consistency of likelihood estimation and stability of tracking. We call a parameter which includes ambiguity "*object position parameter*," a parameter which is independent of object position "*appearance parameter*."

2 Appearance estimation and object tracking

In the following discussion, we propose a stable likelihood estimation for tracking a variable appearance object using the particle filter. We explain how the appearance parameter influences position parameter space in firefly example.

2.1 One dimensional appearance parameter

In the tracking problem that the brightness of target object varies, the appearance parameter space $\mathbb{R}_A$ can be represented as one dimensional space. As the object's appearance and the parameter of appearance have one to one relationship, estimation can use the continuous space. Therefore, we can also treat the problem simply as the particle filter problem in the product space of both of $\mathbb{R}_A$ and position parameter space $\mathbb{R}_P$ as shown in Fig.1. However, each hypothesis differently estimates an appearance parameter $q \in \mathbb{R}_A$. When tracking a blinking firefly in the dark, tracking may fail because hypotheses

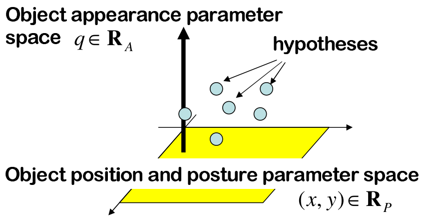

Figure 1. A product space of object's position parameter space and appearance parameter space

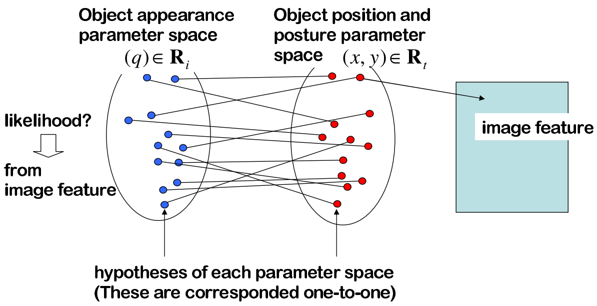

Figure 2. Object state space assumed in this paper: Our method uses a position parameter space for estimating a likelihood of appearance parameter

in the dark position, where target does not exist, become more confident.

For avoiding this tracking failure, we propose a stable likelihood estimation using unique appearance parameter q^*. The target position (x, y) is being tracked estimating q^*. We decompose a target parameter space into two spaces as shown in Fig.2. Estimation of a distribution in $\mathbb{R}_P$ is stabilized by q^*. Then, a likelihood estimation in $\mathbb{R}_A$ is performed by the image feature involving $\mathbb{R}_P$. Our method can keep consistency of the relationship between each space.

Parameters of hypotheses are represented as $\mathbf{s}_t^{(n)} = (x_t^{(n)}, y_t^{(n)}; q_t^{(n)})$ with weights $\pi_t^{(n)}$, where $x_t^{(n)}$ and $y_t^{(n)}$ represent a target position, and $q_t^{(n)}$ represents appearance parameter respectively. The hypothesis represents assumed an appearance parameter $q_t^{(n)}$ for a target position $(x_t^{(n)}, y_t^{(n)})$, where n represents samples' number. The dynamical model for our method is Gaussian noise.

2.2 Position parameter estimation

In time $t = 0$, we can't determine the unique appearance parameter. Then, the first time likelihood is estimated by $q_t^{(n)}$. The likelihood is calculated as

$$\begin{aligned} \pi_{t=0}^{(n)} &= \exp\left(-d_{t=0}^{(n)}/\sigma\right), \\ d_t^{(n)} &= \left|\mathbf{I}_t(x_t^{(n)}, y_t^{(n)}) - \mathbf{A}(q_t^{(n)})\right|^2, \end{aligned} \quad (1)$$

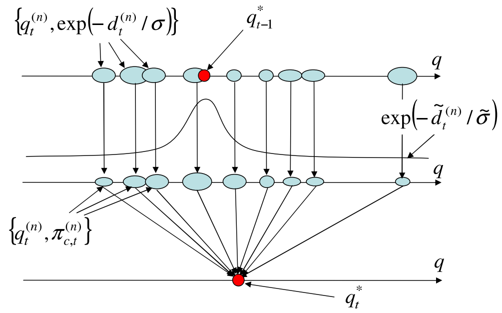

Figure 3. A likelihood estimation in the appearance parameter space

Where, $\mathbf{I}_t(x_t^{(n)}, y_t^{(n)})$ represents a vector of RGB image value[1]. On the target position, $\mathbf{A}(q_t^{(n)})$ represents a vector of RGB value corresponded to appearance parameter. Eq. (1) represents evaluation of matching degree between RGB of image and appearance parameter. In other time $t \neq 0$, we obtain a weight $\pi_t^{(n)}$ as

$$\pi_t^{(n)} = \exp\left(-\left|\mathbf{I}_t(x_t^{(n)}, y_t^{(n)}) - \mathbf{A}(q_{t-1}^*)\right|^2 / \sigma\right), \quad (2)$$

where q_t^* represents the unique appearance parameter which is described following subsection.

2.3 Appearance parameter estimation

As mentioned above, we realize a tracking with q_t^* which is calculated as $q_t^* = \sum_n \pi_{c,t}^{(n)} q_t^{(n)}$. In other words, the expectation with weight $\pi_{c,t}^{(n)}$ of hypotheses $q_t^{(n)}$ represents q_t^*.

Now, we explain how to calculate weight $\pi_{c,t}^{(n)}$ which should indicate correctness of $q_t^{(n)}$. When we calculate $\pi_{c,t}^{(n)}$ in the same way as Eq. (1), appearance parameter $q_t^{(n)}$ representing "a firefly disappears" at the dark position $(x_t^{(n)}, y_t^{(n)})$ is evaluated correct though the firefly is lighting. Because, the assumed target position and the appearance parameter are not correct. For such a reason, we introduce the q_t^* into Eq. (1). Fig.3 represents appearance parameter estimation. Ellipses represent $q_t^{(n)}$, and its size indicates how much $q_t^{(n)}$ fits with image feature. On the middle arrow, the ellipse size indicates weight $\pi_{c,t}^{(n)}$. Around the position of q_{t-1}^*, the ellipses on the middle arrow become large compared with ones on the top arrow. Introducing q_t^* into Eq. (1) gives correct estimation.

However, it is not enough for determining "disappear" or not. Because, almost hypotheses $(x_t^{(n)}, y_t^{(n)})$ don't exist lighting firefly. For solving this problem, a weight at dark positions is set to low value.

For obtaining correct q_t^*, iterative calculation where q_t^* is introduced into Eq. (1) is required because obtained

[1]The reason we don't use grayscale value not RGB value is that we noticed that each RGB values didn't indicate same value in the experiments.

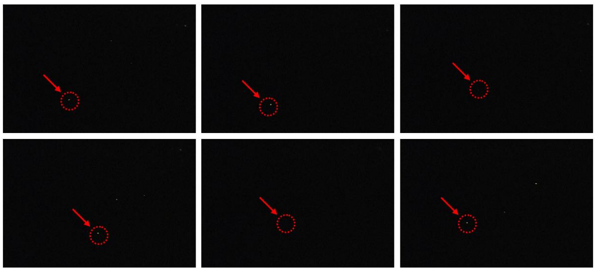

Figure 4. A part of input image sequences. An arrow in this figure represents a target.

value q_t^* is not converged with one-time-only calculation. Then, from the discussion, we propose an estimation of likelihood for $q_t^{(n)}$ and estimation of q_t^* as

$$\begin{aligned} q_{k+1,t}^* &= \sum_n q_t^{(n)} \pi_{k,c,t}^{(n)}, \\ \pi_{k,c,t}^{(n)} &= \frac{\exp\left(-d_t^{(n)}/\tilde{\sigma}\right)\exp\left(-\hat{d}_{k,t}^{(n)}/\hat{\sigma}\right)}{\left|\mathbf{I}_t(x_t^{(n)}, y_t^{(n)}) - \mathbf{I}_b\right| + \epsilon}, \\ \hat{d}_{k,t}^{(n)} &= \left|q_t^{(n)} - q_{k,t}^*\right|^2, \end{aligned} \qquad (3)$$

where $q_{k,t}^*$ denotes k times estimation iteration result of unique appearance parameter, $\pi_{k,c,t}^{(n)}$ denotes weight estimated by likelihood in k time iteration, $\mathbf{I}_b$ denotes background image information (in a case of a problem of a blinking firefly tracking, it represents an RGB equal to zero). In Eq. (3), the result of appearance parameter $q_{k=1,t}^* = q_{t-1}^*, q_t^* = q_{k=K,t}^*$, where K is iteration total.

3 Experimental results

In this paper, we performed tracking experiment by using a movie of a wild firefly. Fig.4 shows input images captured in 30fps. The image frame numbers are frames 89, 178, 267, 356, 445, and 534 from the left top to right bottom. In frame 267 and 445, the firefly disappears. $\mathbf{A}(\cdot)$ is learned beforehand as a quantum in 59 stages. The number of hypotheses is 600. The hypotheses are initialized by user. In this paper, we did three experiments for evaluating effectiveness of proposal method. First, we did the original CONDENSATION [1] experiments (experiments A). Second is an experiment of without iteration of k in Eq. (3); $K = 1$ (experiments B). Third is our method (experiments C).

3.1 The original CONDENSATION result

In the experiment A, we track a firefly using the original CONDENSATION. Fig.5 shows a result of the experiment. Green points represent each hypothesis. An upper right rectangle on Fig.5 represents enlarged part around tracking result. Fig.6 (A) shows a trajectory of the tracking result represented as red line with true value represented as blue lines (when a firefly disappears, we can't get a true value). In this figure, x and y-axis represent an image coordinates, frame axis represents a frame number. From Fig.5 and 6 (A), we can notice that the tracking can't be performed correctly. The reason the tracking fails is that the appearance parameter of a firefly, which is criteria of likelihood estimation, can't be determined correctly and then tracking results in failure.

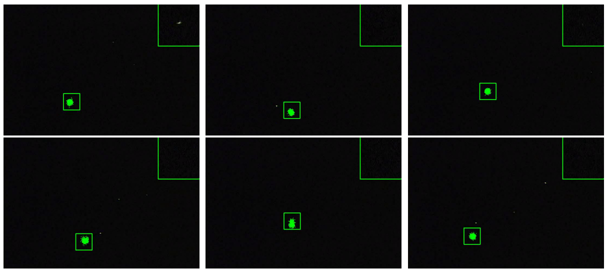

Figure 5. The experiment results A:Points in this figure represent hypotheses.

3.2 Experimental results without iteration

Experimental result B shows a method for determining appearance parameter q_t^* once in a frame. Fig.7 shows the experimental results, and Fig.6 (B) shows a trajectory of the tracking result. In Fig.7, rectangular area on the upper left of image shows the color that corresponds to the estimated appearance parameter. We can notice that the tracking results are more accurate than the experiment A.

Moreover, Fig.8 shows unique appearance parameter estimation result. The horizontal axis indicates time (frame number), vertical axis indicates the appearance parameter estimated in Fig.7. The blue line indicates true value, and the red line indicates estimated value. Mean value of the error margin between true value and estimated value of state is 6.42(degree), standard deviation value is 9.51(degree). Fig.8 shows estimation of unique appearance parameter, the above-mentioned mean value of the error margin and standard deviation value. It is shown that no iterated processing for q_t^* can't estimate correct value.

3.3 Experimental results with iteration

Fig.9 and 6 (C) show results of tracking experiment by our method (including iterated estimation) as well as experiment B.

Fig.10 shows the result of estimated appearance parameter in this experiment. Mean value of the error margin between true value and estimation value of appearance parameter in this experiment is 4.75(degree), standard deviation value is 7.48(degree). These values represent that correct appearance parameter estimation needs an iteration of k. Fig.10 and values show that we realize the estimation of correct appearance parameter. Although, behavior of hypotheses is a little unsteady when firefly disappeared in frame 267 and 445 in Fig.9. However, our method can track the target and estimate an appearance parameter better than any other experiments.

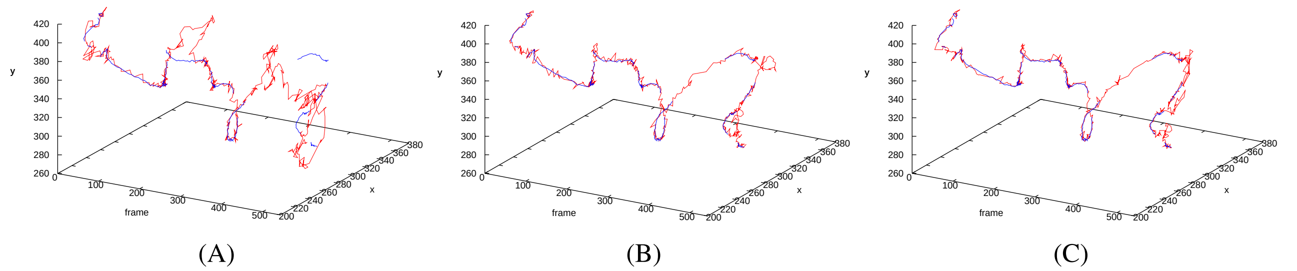

Figure 6. A trajectory of tracked firefly all the experimental results

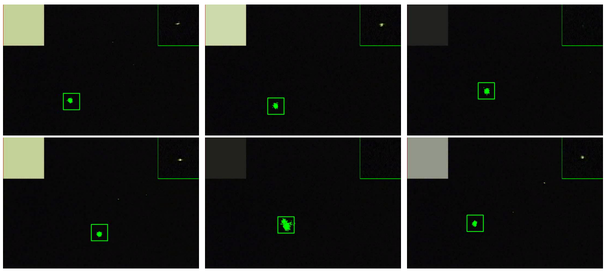

Figure 7. The experiment results B

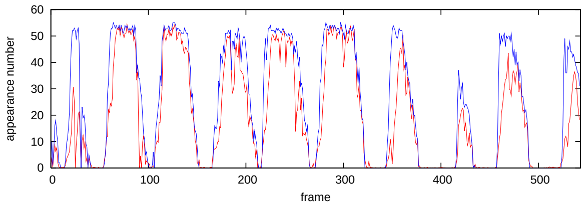

Figure 8. A state estimation results from the experiments B: a horizontal axis indicates frame, a vertical axis indicates q_t^*

4 Conclusion

In this paper, we proposed a stable likelihood estimation for variable appearance object using the particle filter. Our method calculates the unique "*appearance parameter*" of target which is used for likelihood estimation in "*position parameter space*." Hence, we can obtain correct target's "*appearance parameter*" and "*position parameter*."

The experimental results show advantage that can estimate *appearance parameter* "light" or "disappear" and the target position only once, which a generic particle filter can't deal with together. The tracking results that our method can track a target more robust than original CONDENSATION.

In future work, we will do experiments for multiple targets. Synchronizing fireflies in same sequence may occur tracking failure. Multiple distributions are introduced in the problem. We will consider an overlap and a mis-estimation of the distributions.

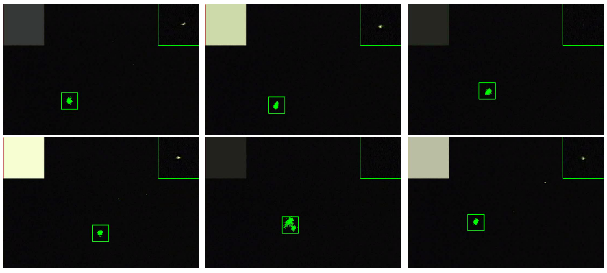

Figure 9. The experiment results C

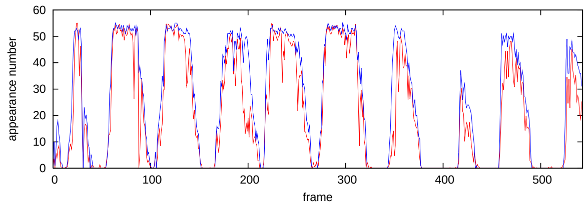

Figure 10. An appearance state estimation results from the experiments C

References

[1] M. Isard and A. Blake. Condensation – conditional density propagation for visual tracking. *International Journal of Conputer Vision*, 29(1):5–28, 1998.

[2] M. Isard and A. Blake. ICondensation: Unifying low-level and high-level tracking in a stochastic framework. In *5th ECCV*, pages 893–908, 1998.

[3] Y. Jin and F. Mokhtarian. Variational particle filter for multi-object tracking. In *ICCV'07*, pages 1–8, 2007.

[4] Z. Khan. MCMC-based particle filtering for tracking a variable number of interacting targets. *IEEE Trans. Pattern Anal. Mach. Intell.*, 27(11):1805–1918, 2005. Member-Tucker Balch and Member-Frank Dellaert.

[5] J. MacCormick and M. Isard. Partitioned sampling, articulated objects, and interface quality hand tracking. In *ECCV*, pages 3–19, 2000.

[6] K. Okuma, A. Taleghani, N. de Freitas, J. Little, and D. Lowe. A boosted particle filter: Multitarget detection and tracking. In *ECCV'04*, pages 28–39, 2004.

[7] P. Perez, J. Vermaak, and A. Blake. Data fusion for visual tracking with particles. *Proceedings of IEEE (issue on State Estimation)*, 92:495–513, Feb 2004.

[8] C. Yang, R. Duraiswami, and L. Davis. Fast multiple object tracking via a hierarchical particle filter. In *ICCV '05*, volume 1, pages 212–219, 2005.