

SENSOR FUSION UNDER UNKNOWN ASSOCIATIONS BY PARTICLE FILTERS WITH CLEVER PROPOSAL

Norikazu Ikoma, Wataru Ito, Masato Kawanishi, and Hiroshi Maeda

Faculty of Engineering, Kyushu Institute of Technology, Fukuoka 804-8550, JAPAN

Email: ikoma@comp.kyutech.ac.jp

ABSTRACT

A new method for sensor fusion under unknown associations among multiple sensors is proposed. Fundamental problem within the sensor fusion situation is huge number of the associations that prohibits to enumerate all the combinations within tractable computational time. Proposed method formulates this situation in a state space model, which is highly nonlinear to deal with the unknown associations, and utilizes particle filters to estimate state of the model. Then we obtain state of the target system as well as the associations through the state estimation. We also propose clever proposal in the framework of particle filters that draws efficient particles in a sense of sub-optimality to minimize the variance of particles' weight. The proposed method is formulated in generic way, so, in principle, it can be applied to various situations for sensor fusion under unknown associations. We show an illustrative example to track sound target in a scene with sensors of two microphones and one camera.

Keywords: *Sensor fusion, state space model, particle filter, nonlinear, clever proposal, target tracking.*

1. INTRODUCTION

Sensor fusion can be defined as a task that reconstructs real world information by using data coming from multiple sensors where it is difficult or impossible to obtain the information by using only single sensor. Most highly organized life, including human, perform this task naturally. Also there are many examples of this task in engineering field, such as 3D reconstruction in computer vision, target detection using multiple sensors (e.g., radar, sonar, and infrared rays), and recognition of environment (e.g., localization, object recognition, and map learning) by mobile robot with multiple sensors' data.

Fundamental problem arises in sensor fusion, which is called association problem. The problem is to determine the unknown correspondence among signals of all sensors. Here it is assumed that the sensor generally detects many signals of the real world objects and correspondence between the signals and the objects is unknown. Then the problem is equivalent to determine the correspondence between the signals of a sensor and the objects, for all sensors. This causes a combinatorial problem, which is factorial order of the number

of objects in the real world when it is assumed that one signal is coming from one object. Furthermore, in actual situation, there are possibilities for missing and false detection of the signals, so the number of combinations will grow very quickly.

There is some approaches to cope with this combinatorial problem by using the idea of Monte Carlo methods. First, with single sensor, determination of the unknown association using particle filters in a context of target tracking is investigated in [4], [5], [6], and [8]. Next, for stereo camera situation as a multiple sensors case, Formulation of the problem in the form with unobserved variable to be able to apply EM algorithm and efficient solution to the problem by Markov Chain Monte Carlo(MCMC) are proposed in [2]. There is also a research for dynamic image situation with stereo camera using particle filters [7].

In this paper, we propose a general idea for sensor fusion under unknown association base on technique of particle filters and clever proposal. The idea is general enough so it includes all the situations mentioned above. We formalize the general situation of sensor fusion by a state space model having the unknown associations in its state. Next, we apply particle filters to estimate state of the model, then we have state of the target system as well as the associations. Here is a key to succeed this estimation task, which is called "clever proposal" in the context of particle filters.

Particle filters, in general, use proposal distribution to draw particles in the state space, and it calculates weight for each particle to adjust the set of particles to the target distribution based on importance sampling idea. Thus how well the particles are drawn controls how well the particle filters work. Clever proposal is a generic term that draws particles trying to minimize variance of the weights. We propose an elaborated proposal as the clever proposal, which is sub-optimal one to trade-off computational tractability and optimality of the minimum variance.

We formulate this idea in the following sections. Firstly we propose a new state space model having unknown associations with its state in general form, then the algorithm of particle filters to estimate the state is shown. Where a clever proposal, which is also a novel idea, is proposed. At the end of this paper, we will show an illustrative example that track a target in the scene, where the target makes sound.

2. MODEL

We propose a new state space model for sensor fusion under unknown associations. The state space model consists of system equation and observation equation. System equation represents dynamics of the target system, and observation equation is a set of equations for all sensors.

System equation is

$$\mathbf{x}_k = \mathbf{f}(\mathbf{x}_{k-1}, \mathbf{v}_k), \quad \mathbf{v}_k \sim q(\cdot), \quad (1)$$

where \mathbf{x}_k is state of target system at discrete time k . The state contains information about N_T objects in the target system. \mathbf{v}_k is i.i.d. random vector having pdf $q(\cdot)$, and it is called "system noise".

Suppose that there are S sensors, and let $s \in \{1, 2, \dots, S\}$ be index variable for the sensors. s -th sensor model is represented by

$$\mathbf{y}_{s,k} = \mathbf{h}_s(\mathbf{x}_k, \mathbf{w}_{s,k}; \Theta_{s,k}), \quad \mathbf{w}_{s,k} \sim r_s(\cdot), \quad (2)$$

where $\mathbf{y}_{s,k}$ is vector of signals detected by s -th sensor, and $\mathbf{w}_{s,k}$ is i.i.d. random vector having pdf $r_s(\cdot)$, and it is called "observation noise".

All noises appeared above, i.e., \mathbf{v}_k and $\mathbf{w}_{s,k}$ for $s = 1, 2, \dots, S$, are assumed to be mutually independent.

Suppose that $\mathbf{y}_{s,k}$ has $N^s(k)$ signals, which is consisting of $N_D^s(k)$ signals coming from the target system and $N_C^s(k)$ signals due to false detection. Thus $N_D^s(k) \leq N_T$ and it follows that $N^s(k) = N_D^s(k) + N_C^s(k)$. $\Theta_{s,k}$ is associations vector having elements as

$$\Theta_{s,k} = \left(\theta_1^s(k), \theta_2^s(k), \dots, \theta_{N^s(k)}^s(k) \right). \quad (3)$$

Each element takes value in $\{0, 1, 2, \dots, N_T\}$, where 0 means the signal is false detection, and other values, $\theta_j^s > 0$, means j -th signal comes from θ_j^s -th object in the target system.

Let $\#(\Theta_{s,k})$ be combination number of the associations of s -th sensor. Then, combination number for all sensors becomes $\prod_{s=1}^S \#(\Theta_{s,k})$. For example, if there is no false detection and missing, and one to one mapping between objects in the target system and detected signals holds, then, $\#(\Theta_{s,k}) = N_T!$. The number becomes much larger for more realistic situations with false detection and missing.

We assume Markov property for time evolution of the associations, such that $\Theta_k \sim p(\cdot | \Theta_{k-1})$, where $\Theta_k = (\Theta_{1,k}, \Theta_{2,k}, \dots, \Theta_{S,k})$. For example, we can assume that the associations, Θ_k , are mutually independent with respect to the sensors, and association of each sensor, $\Theta_{s,k}$, is assumed to be of Markov process with high probability to stay the same state and with small probability to change the state.

We augment the state vector as $\mathbf{z}_k \equiv (\mathbf{x}_k, \Theta_k)$. By denoting $\mathbf{y}_k = (\mathbf{y}_{1,k}, \mathbf{y}_{2,k}, \dots, \mathbf{y}_{S,k})$, we can form an augmented state space model such that

$$\mathbf{z}_k = \mathbf{F}(\mathbf{z}_{k-1}, \mathbf{V}_k) \quad (4)$$

$$\mathbf{y}_k = \mathbf{H}(\mathbf{z}_k, \mathbf{W}_k) \quad (5)$$

with appropriate definitions for system noise vector \mathbf{V}_k and observation noise vector \mathbf{W}_k .

By estimating the augmented state of the state space model (4) and (5), we can obtain the estimation of state of the target system, \mathbf{x}_k , and the associations, Θ_k , simultaneously.

3. PARTICLE FILTERS

3.1. Simple Particle Filter

For the state estimation, we use "particle filters" [3]. Particle filters use many number of weighted particles in the state space to approximate target distribution. The target distribution is conditional distribution of the state given a series of observation such that $p(\mathbf{z}_{1:k} | \mathbf{y}_{1:k})$, where we employ useful notation $\mathbf{z}_{1:k} = (\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_k)$. Particle filters update the conditional distribution, i.e., it is recursive estimation such that

$$p(\mathbf{z}_{1:k} | \mathbf{y}_{1:k}) = p(\mathbf{z}_{1:k-1} | \mathbf{y}_{1:k-1}) \frac{p(\mathbf{y}_k | \mathbf{z}_k) p(\mathbf{z}_k | \mathbf{z}_{k-1})}{p(\mathbf{y}_k | \mathbf{y}_{1:k-1})} \quad (6)$$

by updating the weighted particles. Where $p(\mathbf{y}_k | \mathbf{z}_k)$ and $p(\mathbf{z}_k | \mathbf{z}_{k-1})$ are derived from eq.(5) and (4), respectively.

The algorithm of particle filters is as follows. Suppose that weighted particles which approximate $p(\mathbf{z}_{1:k-1} | \mathbf{y}_{1:k-1})$ are given as $\left\{ \left(\mathbf{z}_{1:k-1}^{(i)}, \omega_{k-1}^{(i)} \right) \right\}_{i=1}^M$, where $\mathbf{z}_{1:k-1}^{(i)}$ is i -th instance of state variable, $\mathbf{z}_{1:k-1}$, which we call "particle", and $\omega_{k-1}^{(i)}$ is corresponding weight, which is non-negative, normalized as $\sum_{i=1}^M \omega_{k-1}^{(i)} = 1$. The algorithm consists of three steps; 1) draw of particles, 2) weight update, and 3) re-sampling. Each step is as follows;

1) Draw of particles:

Draw particles from "proposal distribution", which is denoted by $q(\mathbf{z}_k | \mathbf{z}_{1:k-1}, \mathbf{y}_{1:k})$, such that

$$\tilde{\mathbf{z}}_k^{(i)} \sim q(\cdot | \mathbf{z}_{1:k-1}^{(i)}, \mathbf{y}_{1:k}), \quad i = 1, 2, \dots, M. \quad (7)$$

2) Weight update:

Update weight according to the equation below, which can be derived from eq.(6),

$$\tilde{\omega}_k^{(i)} \propto \omega_{k-1}^{(i)} \frac{p(\mathbf{y}_k | \tilde{\mathbf{z}}_k^{(i)}) p(\tilde{\mathbf{z}}_k^{(i)} | \mathbf{z}_{k-1}^{(i)})}{q(\tilde{\mathbf{z}}_k^{(i)} | \mathbf{z}_{1:k-1}^{(i)}, \mathbf{y}_{1:k})}, \quad i = 1, 2, \dots, M, \quad (8)$$

note that $\tilde{\omega}_k^{(i)}$ is normalized. Now, we let $\tilde{\mathbf{z}}_{1:k}^{(i)} \equiv \left(\mathbf{z}_{1:k-1}^{(i)}, \tilde{\mathbf{z}}_k^{(i)} \right)$, then we have set of weighted particles, $\left\{ \left(\tilde{\mathbf{z}}_{1:k}^{(i)}, \tilde{\omega}_k^{(i)} \right) \right\}_{i=1}^M$, as an approximation of $p(\mathbf{z}_{1:k} | \mathbf{y}_{1:k})$.

3) Re-sampling:

Sample particles from $\left\{ \left(\tilde{\mathbf{z}}_{1:k}^{(i)}, \tilde{\omega}_k^{(i)} \right) \right\}_{i=1}^M$, where sampling proceeds with probability $\tilde{\omega}_k^{(i)}$ to draw particle $\tilde{\mathbf{z}}_{1:k}^{(i)}$. Denote the drawn particle by $\mathbf{z}_{1:k}^{(i)}$, and let

$\omega_k^{(i)} := 1/M$, then we have $\left\{ \left(\mathbf{z}_{1:k}^{(i)}, \omega_k^{(i)} \right) \right\}_{i=1}^M$ as another approximation of $p(\mathbf{z}_{1:k} | \mathbf{y}_{1:k})$.

3.2. Rao-Blackwellized Particle Filter

To have efficient performance of particle filters, we further introduce an elaborated idea called "Rao-Blackwellization (RB)" [1], which decomposes the state into analytical part and particle approximation part as

$$p(\mathbf{z}_{1:k} | \mathbf{y}_{1:k}) = p(\mathbf{x}_{1:k} | \mathbf{y}_{1:k}, \Theta_{1:k}) p(\Theta_{1:k} | \mathbf{y}_{1:k}), \quad (9)$$

where $p(\Theta_{1:k} | \mathbf{y}_{1:k})$ is approximated by weighted particles, and $p(\mathbf{x}_{1:k} | \Theta_{1:k}, \mathbf{y}_{1:k})$ is calculated based on Kalman filter.

Through the same derivation of eq.(6), we have

$$\begin{aligned} p(\Theta_{1:k} | \mathbf{y}_{1:k}) &= p(\Theta_{1:k-1} | \mathbf{y}_{1:k-1}) \\ &\times \frac{p(\mathbf{y}_k | \mathbf{y}_{1:k-1}, \Theta_{1:k-1}) p(\Theta_k | \Theta_{k-1})}{p(\mathbf{y}_k | \mathbf{y}_{1:k-1})}. \end{aligned} \quad (10)$$

Then, the algorithm becomes as follows, here we change the notation of weight from ω_k to ν_k .

1) Draw of particles (RB):

Draw particles, i.e. associations, from proposal π as,

$$\tilde{\Theta}_k^{(i)} \sim \pi(\cdot | \Theta_{1:k-1}^{(i)}, \mathbf{y}_{1:k}), \quad i = 1, 2, \dots, M. \quad (11)$$

2) Kalman update (RB):

In order to evaluate $p(\mathbf{y}_k | \mathbf{y}_{1:k-1}, \tilde{\Theta}_{1:k}^{(i)})$, which will be required for weight update step, we perform Kalman filter as follows. Suppose that filtering distribution of analytical part at time $k-1$, $p(\mathbf{x}_{k-1} | \mathbf{y}_{1:k-1}, \Theta_{1:k-1}^{(i)})$, is given as Gaussian distribution with mean vector $\bar{\mathbf{x}}_{k-1|k-1}^{(i)}$ and covariance matrix $\mathbf{V}_{k-1|k-1}^{(i)}$, for $i = 1, 2, \dots, M$. By applying one-step-ahead prediction step, we have Gaussian distribution $p(\mathbf{x}_k | \mathbf{y}_{1:k-1}, \Theta_{1:k}^{(i)})$ parametrized by $\bar{\mathbf{x}}_{k|k-1}^{(i)}$ and $\mathbf{V}_{k|k-1}^{(i)}$, for $i = 1, 2, \dots, M$. Next, from the one-step-ahead prediction distribution, we calculate prediction distribution of observation, $p(\mathbf{y}_k | \mathbf{y}_{1:k}, \Theta_{1:k}^{(i)})$, which is also Gaussian with mean vector $\bar{\mathbf{y}}_{k|k-1}^{(i)}$ and covariance matrix $\Sigma_{k|k-1}^{(i)}$. Finally, applying filtering step, we have Gaussian distribution $p(\mathbf{x}_k | \mathbf{y}_{1:k}, \Theta_{1:k}^{(i)})$ with parameters $\bar{\mathbf{x}}_{k|k}^{(i)}$ and $\mathbf{V}_{k|k}^{(i)}$, for $i = 1, 2, \dots, M$. Note that if the model of analytical part is nonlinear, then we use extended Kalman filter, which uses linearization of the nonlinear equations, in the above steps.

3) Weight update (RB):

For $i = 1, 2, \dots, M$, update weight by

$$\tilde{\nu}_k^{(i)} \propto \nu_{k-1}^{(i)} \frac{\Phi(\mathbf{y}_k; \bar{\mathbf{y}}_{k|k-1}^{(i)}, \Sigma_{k|k-1}^{(i)}) p(\tilde{\Theta}_k^{(i)} | \Theta_{k-1}^{(i)})}{\pi(\tilde{\Theta}_k^{(i)} | \Theta_{1:k-1}^{(i)}, \mathbf{y}_{1:k})}, \quad (12)$$

where $\Phi(\mathbf{y}; \bar{\mathbf{y}}, \Sigma)$ is pdf of Gaussian distribution with mean \mathbf{y} and covariance Σ . And let $\tilde{\Theta}_{1:k}^{(i)} \equiv (\Theta_{1:k-1}^{(i)}, \tilde{\Theta}_k^{(i)})$.

4) Re-sampling (RB):

In RB, we sample particles with analytical part, i.e., mean vector $\bar{\mathbf{x}}_{k|k}^{(i)}$ and covariance matrix $\mathbf{V}_{k|k}^{(i)}$. So we perform resampling from $\left\{ \left(\left[\tilde{\Theta}_{1:k}^{(i)}, \bar{\mathbf{x}}_{k|k}^{(i)}, \mathbf{V}_{k|k}^{(i)} \right], \tilde{\nu}_k^{(i)} \right) \right\}_{i=1}^M$ in similar way as simple particle filter.

3.3. Clever Proposal

Here is a choice of the proposal distribution. It is a key to make particle filters be efficient. The optimal proposal has been proposed in the literatures as

$$\begin{aligned} \pi(\Theta_k | \Theta_{1:k-1}^{(i)}, \mathbf{y}_{1:k}) &= p(\Theta_k | \Theta_{1:k-1}^{(i)}, \mathbf{y}_{1:k}) \\ &= \frac{p(\mathbf{y}_k | \mathbf{y}_{1:k-1}, \Theta_{1:k-1}^{(i)}, \Theta_k) p(\Theta_k | \Theta_{k-1}^{(i)})}{p(\mathbf{y}_k | \mathbf{y}_{1:k-1}, \Theta_{1:k-1}^{(i)})}, \end{aligned} \quad (13)$$

where the optimality is in a sense of minimum variance of the weight over the set of particles. Note that denominator in eq.(13) is summation of numerator with respect to all combinations of the associations, Θ_k . Thus, in our model, the optimal proposal is not tractable due to huge number of combinations of the associations. To circumvent this difficulty, we propose a novel sub-optimal proposal that trade off between optimality and tractability as follows.

First, we define a proposal that can be decomposed into each sensor as

$$\pi(\Theta_k | \Theta_{1:k-1}^{(i)}, \mathbf{y}_{1:k}) \equiv \prod_{s=1}^S \pi(\Theta_{s,k} | \Theta_{s,1:k-1}^{(i)}, \mathbf{y}_{s,1:k}) \quad (14)$$

where $\mathbf{y}_{s,1:k} \equiv (\mathbf{y}_{s,1}, \mathbf{y}_{s,2}, \dots, \mathbf{y}_{s,k})$. Second, define a proposal for each sensor that can be decomposed into each signal as

$$\begin{aligned} \pi(\Theta_{s,k} | \Theta_{s,1:k-1}^{(i)}, \mathbf{y}_{s,1:k}) &\equiv \pi(\theta_1^s(k) | \Theta_{s,1:k-1}^{(i)}, \mathbf{y}_1^s(k)) \\ &\times \pi(\theta_2^s(k) | \theta_1^s(k), \Theta_{s,1:k-1}^{(i)}, \mathbf{y}_2^s(k)) \\ &\times \pi(\theta_3^s(k) | \theta_1^s(k), \theta_2^s(k), \Theta_{s,1:k-1}^{(i)}, \mathbf{y}_3^s(k)) \\ &\dots \times \pi(\theta_{N^s(k)}^s(k) | \theta_1^s(k), \dots, \theta_{N^s(k)-1}^s(k), \\ &\quad \Theta_{s,1:k-1}^{(i)}, \mathbf{y}_{N^s(k)}^s(k)). \end{aligned} \quad (15)$$

Third and finally, each component in eq.(15) is defined so as to be proportional to a product of roughly approximated likelihood and transition probability such that

$$\begin{aligned} \pi(\theta_j^s(k) | \theta_1^s(k), \dots, \theta_{j-1}^s(k), \Theta_{s,1:k-1}^{(i)}, \mathbf{y}_j^s(k)) &\propto \Phi(\mathbf{y}_j^s(k); \bar{\mathbf{y}}_j^s(k), \hat{\Sigma}_{\theta_j^s(k)}^s) \\ &\times p(\theta_j^s(k) | \theta_1^s(k), \theta_2^s(k), \dots, \theta_{j-1}^s(k), \theta_j^s(k) (k-1)), \end{aligned} \quad (16)$$

where $\bar{\mathbf{y}}_j^s(k)$ is j -th signal element of prediction mean of observation, $\bar{\mathbf{y}}_{k|k-1}^{(i)}$, and $\hat{\Sigma}_{\theta_j^s(k)}^s$ is time constant covariance: it takes specified value for normal situation when $\theta_j^s(k) > 0$, otherwise it takes value for false detection.

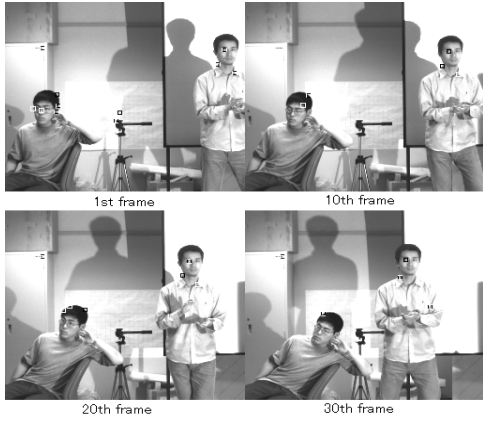


Fig. 1. Image sequence.

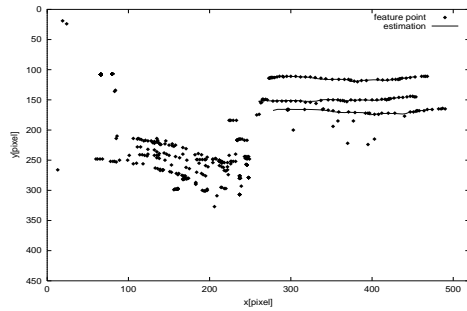


Fig. 2. Tracking result of sound target.

4. EXAMPLE

As an example of sensor fusion, we demonstrate sound target tracking using one camera and two microphones with a scene peoples are in motion. Some frames of the captured image sequence are shown in Fig.1, where a man, right hand side, makes sound by clapping his hand, and the other is moving without sound.

For each frame of the image sequence, feature points are extracted by corner detector based on a local feature of the image as plotted in Fig.1. From recorded sound by the two microphones, we calculate direction of sound as plotted in Fig.3.

By applying the proposed method, we have estimated state of the sound target, which consists of position of each feature point and velocity of the target object. Estimated result of positions are shown in Fig.2 with observed data, positions of feature points. By looking at the result, we can see that smooth trajectory of the sound target is obtained.

Estimated result of the association is as follows. Beginning part of the image frames starting from the time when first sound signal occurred is shown in table 1. In this result, all associations are correctly estimated within the table, and mostly correct throughout the image frames.

5. CONCLUSION

We have proposed a general model for sensor fusion in a form of nonlinear state space model having unknown

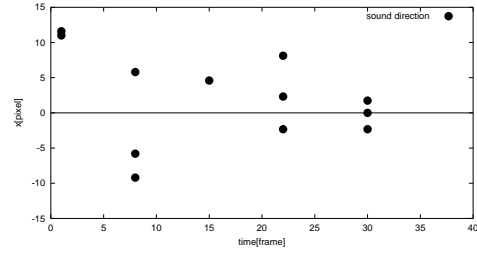


Fig. 3. Sound direction.

Table 1. Estimated associations.

k	feature points' associations										sound assoc.		
7	0	0	1	2	3	0						1	0
8	0	0	1	2	3	0	0						
9	0	1	2	3	0	0							
10	1	2	3	0	0	0							
11	0	1	2	3	0	0	0	0	0	0	0	0	0
12	0	1	2	3	0	0	0	0	0	0	0	0	0
13	0	0	1	2	0	0	0	0	0	0	0	0	0
14	0	0	1	2	3	0	0	0	0	0	0	0	0
15	0	0	1	2	3	0	0	0	0	0	0	0	0
16	0	0	1	2	3	0	0	0	0	0	0	0	0

associations in its state. Particle filters with clever proposal can effectively estimate the state, which consists of state of the target system and the associations. An illustrative example to track sound target shows a reasonable performance to obtain smooth trajectory of sound target as well as the associations.

Acknowledgement: This study was carried out under the ISM Cooperative Research Program (2004-ISM-CPR-2020).

6. REFERENCES

- [1] G.Casella and C.P.Robert: "Rao-Blackwellisation of Sampling Schemes", *Biometrika*, **83**, 1, pp.81-94, 1996.
- [2] F.Dellaert, S.Seitz, and C.E.Thorpe: "EM, MCMC, and Chain Flipping for Structure from Motion with Unknown Correspondence", *Machine Learning*, **50**, pp.45-71, 2003.
- [3] A.Doucet, J.F.G.de Freitas, and N.J.Gordon (eds): *Sequential Monte Carlo Methods in Practice*, New York, Springer, 2001.
- [4] N.J.Gordon and D.J.Salmond: "Bayesian target tracking after group pattern distortion", *Proc. of the SPIE*, **3163**, pp.238-248(1997).
- [5] N.J.Gordon and A.Doucet: "Sequential Monte Carlo for Manoeuvring Target Tracking in Clutter", *Proc. of the SPIE*, **3809**, pp.493-500, 1999.
- [6] N.Ikoma and S.Godsill: "Extended object tracking with unknown association, missing observations, and clutter using particle filters", *Proc. of the 2003 IEEE Workshop on Statistical Signal Processing*, pp.485-488, 2003.
- [7] N.Ikoma, W.Ito, T.Irie, and H.Maeda: "3D Reconstruction from Stereo Camera Dynamic Image based on Particle Filter", *Proc. of International Conference on Intelligent Technologies 2003*, pp.394-403, 2003.
- [8] J.Vermaak, N.Ikoma, and S.Godsill: "Extended Object Tracking using Particle Techniques", *Proc. of the 2004 IEEE Aerospace Conference*, #1015, 2004.