M2SIR: A MULTI MODAL SEQUENTIAL IMPORTANCE RESAMPLING ALGORITHM FOR PARTICLE FILTERS

Thierry Chateau
LASMEA, Clermont-Fd, France
Thierry.CHATEAU@univ-bpclermont.fr

Yann Goyat
LCPC, Nantes, France
yann.goyat@lcpc.fr

Laurent Trassoudaine,
LASMEA, Clermont-Fd, France
Laurent.TRASSOUDAINE@univ-bpclermont.fr

ABSTRACT

We present a multi modal sequential importance resampling particle filter algorithm for object tracking. We consider a hidden state sequence linked to several observation sequences given by different sensors. In a particle filter based framework, each sensor provides a likelihood (weight) associated to each particle and simple rules are applied to merge the different weights such as addition or product. We propose an original algorithm based on likelihood ratios to merge the observations within the sampling step. The algorithm is compared with classic fusion operations on toy examples. Moreover, we show that the method gives satisfactory results on a real vehicle tracking application.

1. INTRODUCTION

Object Tracking is a necessary task for many applications like video surveillance, robotics or Human Machine Applications and many algorithms have been proposed to handle this task. Moreover, tracking an object from several observations is still challenging because sensors deliver correct measures only for nominal conditions (for example the observation of a camera can be identified for a bright and non smoggy day and illumination conditions may change during the tracking process). It results that the fusion process must handle with different probability density functions (pdf) provided by several sensors.This fusion is then a challenging operation because several operators (addition, multiplication, mean, median,...) can be used, which advantages and drawbacks. We propose an original algorithm based on likelihood ratios to merge the observations within the sampling step of a particle filter.

Particle filtering in a visual tracking context has been introduced in [1]. Then, extention to tracking with data fusion has been developed in [2] (a wide bibliography is proposed) in an audiovisual context: different cues are modeled by data likelihood function and intermittent cues are handled. Particle filtering is now very popular for data fusion within a tracking context. Klein [3] propose to introduce belief functions and different combination rules to access particles weight for road obstacle tracking. In a multiple cameras tracking context, Wang [4] propose to adapt the importance sampling method to the data quality. For a similar application, Du [5] propose to combine an independent transition kernel with a booster function to get a mixture function.

The next section details the multi modal sequential importance resampling particle filter method and the associated algorithm. Section three presents the experiments achieved on both synthetic and real data to illustrate the behavior of the proposed algorithm.

2. THE METHOD

2.1. Particle Filter for Several Sources

Particle filtering [6, 1] is a stochastic temporal filter based on the estimation of the a posteriori probability density $p(\mathbf{X}_t|\mathbf{Z}_{0:t})$ of state $\mathbf{X}_t$ conditioned by the historical sequence of observation $\mathbf{Z}_{0:t}$, at time t, by a set of N weighted particles $\{(\mathbf{X}_t^n, \pi_t^n)\}_{n=1}^N$ with their associated weights. The resulting posterior is then approximated by:

$$p(\mathbf{X}_t|\mathbf{Z}_{0:t}) \approx \sum_{n=1}^{N} \pi_t^n.\delta(\mathbf{X}_t - \mathbf{X}_t^n) \tag{1}$$

When the observation is provided by several sources, the likelihood associated to each particle results to the fusion of several weights. This fusion is then a challenging operation because several operators can be used, with advantages and drawbacks. We are proposing to merge observations intrinsically during the re-sampling step of the particle filter .The resulting algorithm (see Algorithm 1) is a variant of the CONDENSATION algorithm [1]. The difference between this algorithm and CONDENSATION is that the weight associated to each particle is a weight vector (composed of weights generated from observations of each source) and that the sampling step is provided by the M2SIR algorithm developed in the following section.

Algorithm 1 CONDENSATION in the multi-source case

Init : particles $\{(\mathbf{X}_0^{\prime n}, \mathbf{1}/N)\}_{n=1}^N$ according to the initial distribution $\mathbf{X}_0$
for $t = 1, ..., T_{end}$ **do**
 Prediction : generation of $\{(\mathbf{X}_t^n, \mathbf{1}/N)\}_{n=1}^N$ from $p(\mathbf{X}_t|\mathbf{X}_{t-1} = \mathbf{X}_{t-1}^{\prime n})$
 Observation : estimation of the weight vector according to the various sources $\{(\mathbf{X}_t^n, \boldsymbol{\pi}_t^n)\}_{n=1}^N$ with $\boldsymbol{\pi}_t^n \propto \mathbf{p}(\mathbf{Z}_t|\mathbf{X}_t = \mathbf{X}_t^n)$
 Sampling : build $\{(\mathbf{X}_{t-1}^{\prime n}, \mathbf{1}/N)\}_{n=1}^N$ from $\{(\mathbf{X}_0^n, \boldsymbol{\pi}_0^n)\}_{n=1}^N$ using M2SIR)
 Estimation : $\hat{\mathbf{X}}_t \doteq \frac{1}{N}\sum_{n=1}^N \mathbf{X}_t^n$
end for
Output : The set of estimated states during the video sequence $\{\hat{\mathbf{X}}_t\}_{t=1,...,T_{end}}$

2.2. M2SIR Algorithm

We consider the estimation of the posterior $p(\mathbf{X}_t|\mathbf{Z}_{0:t})$ at time t, by a set of N particles $\{(\mathbf{X}_t^n, \boldsymbol{\pi}_t^n)\}_{n=1}^N$ with N associated weight vector $\boldsymbol{\pi}_t^n$. The weight vector, off size M given by the number of observations (sources), is composed by the weights related to the sources. For readability, we omit the temporal index t in the following equations. The aim of the proposed multi modal sequential importance resampling algorithm (M2SIR) is to generate a new particle with a three step approach, illustrated in Fig. 1 in the case of three sources

1. M samples (one for each source) are drawn using an Importance Sampling strategy. The resulting output of the step is a set of M candidate samples and their associated weight vector: $\{\mathbf{X}^{(i)}, \boldsymbol{\pi}^{(i)}\}_{i=1,...,M}$

2. A likelihood ratio vector $\mathbf{r}$ off size M is then built from likelihood ratios estimated for each candidate sample. (see below for more details).

3. The selected candidate sample is finally given by an importance sampling strategy operated on a normalized likelihood ratio vector.

The M likelihood ratios used in step two, called r_i ($i = 1, .., M$) are computed by:

$$r_i \doteq \prod_{j=1}^{M}\prod_{k=1}^{M}\left(\frac{\pi_j^i}{\pi_j^k}\right) \tag{2}$$

Equation 2 can be written in a simplest way using log ratio:

$$lr_i = \sum_{j=1}^{M}\sum_{k=1}^{M}\left[\log\left(\pi_j^i\right) - \log\left(\pi_j^k\right)\right] \tag{3}$$

where lr_i denotes the log of r_i. Finally, lr_i is given by:

$$lr_i = M\sum_{j=1}^{M}\left[\log\left(\pi_j^i\right) - \frac{1}{M}\sum_{k=1}^{M}\log\left(\pi_j^k\right)\right] \tag{4}$$

If $\mathbf{l}r \doteq (lr_1, ..., Lr_M)^T$ denotes the vector composed by the log ratios lr_i and $\mathbf{l}\boldsymbol{\pi}^{\mathbf{k}} \doteq (\log \pi_1^k, ..., \log \pi_M^k)^T$ denotes the vector composed by the log of π_j^k, $\mathbf{l}r$ can be written:

$$\mathbf{l}r \doteq M\begin{pmatrix} \mathbf{1}_{(1\times M)}\left(\mathbf{l}\boldsymbol{\pi}^1 - \frac{1}{M}\sum_{k=1}^{M}\mathbf{l}\boldsymbol{\pi}^k\right) \\ \mathbf{1}_{(1\times M)}\left(\mathbf{l}\boldsymbol{\pi}^2 - \frac{1}{M}\sum_{k=1}^{M}\mathbf{l}\boldsymbol{\pi}^k\right) \\ \dots \\ \mathbf{1}_{(1\times M)}\left(\mathbf{l}\boldsymbol{\pi}^M - \frac{1}{M}\sum_{k=1}^{M}\mathbf{l}\boldsymbol{\pi}^k\right) \end{pmatrix} \tag{5}$$

with $\mathbf{1}_{(1\times M)}$ a matrix off size one line and M columns filled by ones. if $\mathbf{C}_\pi \doteq \frac{1}{M}\sum_{k=1}^{M}\mathbf{l}\boldsymbol{\pi}^k$, $\mathbf{l}r$ can be written:

$$\mathbf{l}r = M\begin{pmatrix} \mathbf{1}_{(1\times M)}\left(\mathbf{l}_{\boldsymbol{\pi}^1} - \mathbf{C}_\pi\right) \\ \mathbf{1}_{(1\times M)}\left(\mathbf{l}_{\boldsymbol{\pi}^2} - \mathbf{C}_\pi\right) \\ \dots \\ \mathbf{1}_{(1\times M)}\left(\mathbf{l}_{\boldsymbol{\pi}^M} - \mathbf{C}_\pi\right) \end{pmatrix} \tag{6}$$

$\mathbf{l}r$ represents an unnormalized log. weight vector and the final normalized weight vector is given by:

$$\mathbf{c} \doteq C_c.\exp\left(\mathbf{l}r\right) \tag{7}$$

where $C_c \doteq \mathbf{1}_{(1\times M)}\mathbf{l}r$. $\mathbf{r}$ is then used in step three to select a sample for the M candidates with a importance sampling strategy.

Algorithm 2 M2SIR

Input : Particle set and associated weight vector $\{\mathbf{X}^{(i)}, \boldsymbol{\pi}^i\}_{i=1,...,N}$, M sources
for $n = 1$ to N **do**
 - Choose M candidate particles on the basis of $\{\mathbf{X}^{(i)}, \boldsymbol{\pi}^{(i)}\}_{i=1,...,N}$ and build $\{\mathbf{X}^{*(j)}, \boldsymbol{\pi}^{*(j)}\}_{j=1,...,M}$ where $\mathbf{X}^{*(j)}$ is derived from an *importance sampling* drawn on source j weights;
 - Calculate vector $\mathbf{l}_r$ based on Equation 6, and then calculate confidence vector $\mathbf{c} \doteq C_c.\exp\left(\mathbf{l}_r\right)$
 - Select the designated particle $\mathbf{X}^{e(n)}$ from among the candidate particles by proceeding with an *importance sampling* drawing.
end for
Output : Particle set $\{\mathbf{X}^{e(i)}\}_{i=1,...,N}$ composed of the selected particles.

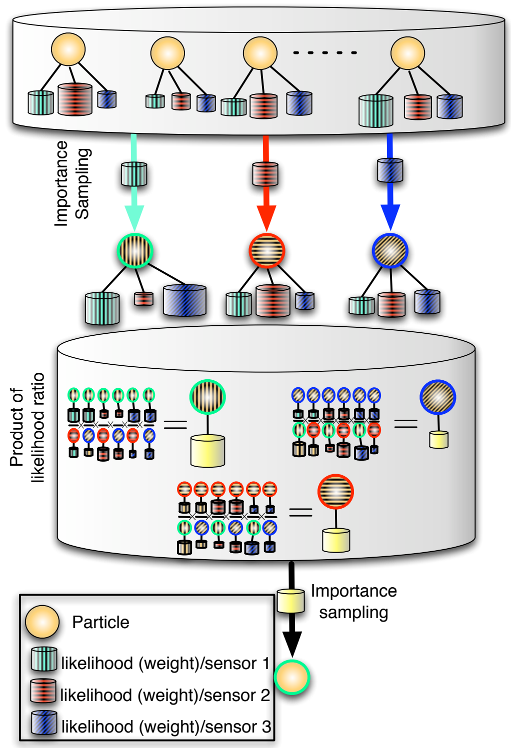

Fig. 1. synoptic of the M2SIR algorithm in the case of three sources: 1)Three particles are drawn using importance sampling (one for each sensor weight distribution). 2) Likelihood ratio are then computed for the three particles. 3) The final particle is drawn with importance sampling from the three ratios.

3. EXPERIMENTS

To validate the method, experiments have been achieved on both synthetic and real data. We first show the behavior of the sampling process for several toy examples generated using simulation. A second set of experiments illustrates the method for a real multi-sensor tracking application.

3.1. Synthetic Data

The aim of this experiment is to compare the behavior of the proposed algorithm with an importance sampling strategy applied to the sum (called SSIR) or the product (called PSIR) of weights (pdf) provided by three sensors. The first example (cf. fig 2) illustrates the behavior of the algorithm when two sensors gives dissonant pdf while the third is blind (uniform pdf). In this example, both the SSIR and M2SIR methods give a resulting pdf reporting the two modes present in the pdf of sensors two and three. The PSIR method provides a third ghost mode in between modes of sensors 2 and 3. The second example (cf. fig 3) compares the three fusion approaches in the case of one blind sensor while the two other ones provide the same pdf. in this case, the SSIR method process a noisy pdf resulting to the blind sensor. Both PSIR and M2SIR gives the same pdf, decreasing the variance of sensors 2 and 3.

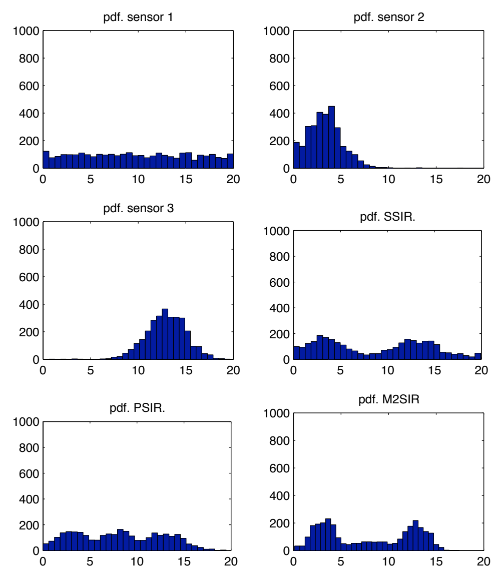

Fig. 2. Illustration of multi-source sampling algorithm for a three sensor fusion step. The pdf provided from sensor one is blind (follows a uniform law) while the pfd provided by sensors two and three have dissonant pdf. SSIR and PSIR are computed using a importance sampling strategy applied respectively to the sum (product) of particles weight.

3.2. Real Data: Application to Vehicle Tracking

The method has been used for vehicle tracking for a static sensor composed by a camera and a laser rangefinder (cf. figure 4). Details of the method can be find here [7]. In order to estimate the precision of the algorithms, ground truth has been acquired using a RTKGPS[1]. A set of twenty sequences at different velocities and under different illumination conditions has been acquired with the associated RTKGPS trajectories. A calibration step gives the homography between the image plane and and GPS ground plane such as an average error can be computed in centimeters into the GPS reference frame. Table 1 shows the estimated precision provided by three fusion strategies: PSIR, SSIR and M2SIR. Results provided by the M2SIR is slighty better than SSIR and PSIR. An other set of twenty sequences has been acquire with a unplugged sensor with provides constant mea-

[1] Real time Kinematics GPS with a precision up to 1cm

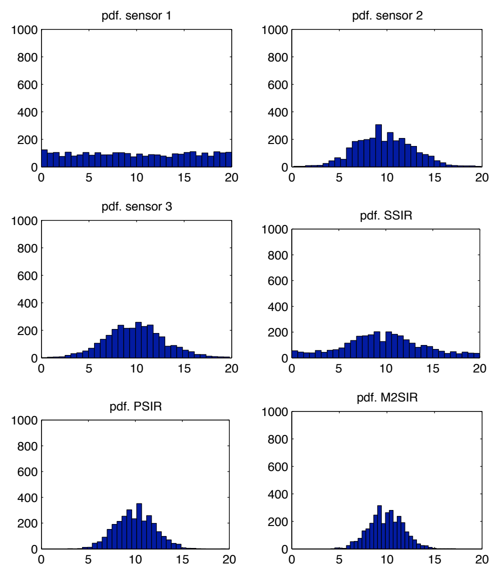

Fig. 3. Illustration of multi-source sampling algorithm for a three sensor fusion step. pdf provided from sensor one is blind (follows a uniform law) while pfd provided by sensors two and three are the same (Gaussian law). pdf. SSIR and PSIR are computed using a importance sampling strategy applied respectively to the sum (product) of particles weight.

sures. Table 2 shows the estimated precision provided by three fusion strategies. The SSIR fusion strategy provides a poor precision comparing to PSIR and M2SIR.

	SSIR	**PSIR**	**M2SIR**
mean/cm	0.16	0.16	0.15
std.	0.10	0.11	0.10

Table 1. Trajectories error for three fusion strategies.

4. CONCLUSION

We have presented a multi modal sequential importance resampling particle filter algorithm for object tracking. The method, based on likelihood ratios, can be used easily within a particle filter algorithm. Experiments show that the method deals efficiently with both blind and dissonant sensors. Moreover, the method has been tested into a real tracking application and gives good results. However, further tests have to be done in order to demonstrate that M2SIR outperform classic fusion operators like product or sum.

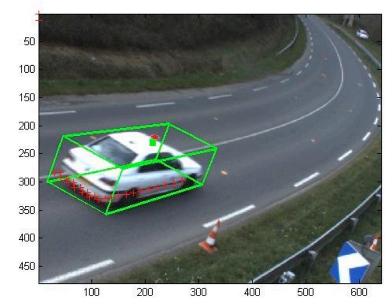

Fig. 4. Ground truth = GPS (red/dark). Estimated trajectory = virtual GPS antenna on the tracking cube (green/clear).

	SSIR	**PSIR**	**M2SIR**
mean/cm	0.22	0.12	0.12
std.	0.12	0.07	0.07

Table 2. Trajectories error for three fusion strategies (one sensor has been unplugged to provide wrong data (constant).

5. REFERENCES

[1] M. Isard and A. Blake, "Condensation – conditional density propagation for visual tracking," *IJCV : International Journal of Computer Vision*, vol. 29, no. 1, pp. 5–28, 1998.

[2] J. Vermaak P. Pérez and A. Blake, "Data fusion for visual tracking with particles," *Proceedings of the IEEE*, vol. 92, no. 2, pp. 495–513, 2004.

[3] J. Klein, C. Lecomte, and P. Miche, "Preceding car tracking using belief functions and a particle filter," in *ICPR08*, 2008, pp. 1–4.

[4] J.K. Wu Y.D. Wang and A. Kassim, "Adaptive particle filter for data fusion of multiple cameras," *The Journal of VLSI Signal Processing*, vol. 49, no. 3, pp. 363–376, 2007.

[5] W. Du, Y. Maret, and J. Piater, "Multi-camera people tracking by collaborative particle filters and principal axis-based integration," in *ACCV*, 2007, pp. 365–374.

[6] S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for on-line non-linear/non-gaussian bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, Feb. 2002.

[7] Y. Goyat, T. Chateau, and L. Trassoudaine, "Tracking of vehicle trajectory by combining a camera and a laser rangefinder," *Springer MVA : Machine Vision and Application*, vol. to appear, 2009.