

Multiple Model Particle Filter

Nadia OUDJANE, Christian MUSSO

ONERA/DTIM/MCT

29, av. de la division Leclerc, 92320 Châtillon, France

oudjane@onera.fr, musso@onera.fr

Résumé – Dans cet article, on s'intéresse au problème du filtrage pour les systèmes à modèles dynamiques multiples appliqué à la poursuite de cible manœuvrante. On choisit l'approche "particulaire" (type Monte Carlo), récemment proposée pour les problèmes de filtrage, car elle permet d'approcher, de manière non paramétrique et indépendante de la dimension la loi conditionnelle de l'état sachant les mesures, même dans les cas non linéaires ou non gaussiens. Le nouveau type de filtre particulaire, présenté ici, (Multiple Model Particle Filter, MMPF) est validé par son bon comportement lors des simulations.

Abstract – This article deals with the filtering problem for multiple dynamical models systems applied to the tracking of a maneuvering target. We chose the particle approach recently proposed for the nonlinear filtering problem. Particle methods can give a non parametric approximation to the signal's conditional distribution even in nonlinear or non Gaussian cases, without depending on the state space dimension. We present a new version of particle filter, called the Multiple Model Particle Filter (MMPF), which is validated by its good behavior during simulations.

1 Introduction

Dans ce papier, nous présentons une nouvelle approche du problème de filtrage, pour les systèmes à modèles dynamiques multiples. Le but est d'estimer, à l'instant courant, t discret, la valeur d'un processus aléatoire,

$$X_t = F_{\theta_t}(X_{t-1}, W_t), \quad (X_t, W_t) \in \mathbb{R}^d \times \mathbb{R}^d \quad (1)$$

indirectement lié à un processus d'observations dont on connaît une réalisation,

$$Y_t = H_{\theta_t}(X_t) + V_t, \quad V_t \perp\!\!\!\perp X_t \quad V_t \in \mathbb{R}^q \quad (2)$$

$(W_t)_{t \geq 1}$ et $(V_t)_{t \geq 1}$ sont des bruits indépendants (pas nécessairement gaussiens) et $(\theta_t)_{t \geq 1}$ est une chaîne de Markov homogène, prenant ses valeurs dans $\{1, \dots, M\}$ et déterminant, à chaque instant, le modèle de dynamique (F_1, \dots, F_M) et d'observation (H_1, \dots, H_M) du système. On suppose que θ_t est indépendant de X_{t-1} sachant θ_{t-1} . Les distributions de X_0 et de θ_0 sont supposées connues ainsi que les lois de transitions, $\mathcal{L}(\theta_t|\theta_{t-1})$ et $\mathcal{L}(X_t|\theta_t, X_{t-1})$. On verra qu'en filtrage particulaire, il suffit de savoir simuler ces lois.

À l'heure actuelle, l'Interacting Multiple Model (IMM) ([1]) est l'outil le plus communément utilisé pour ce type de problème. Le principe de ce filtre consiste à faire interagir autant de Filtres de Kalman Etendus (EKF) que de modèles. C'est pourquoi, dans le cas général du filtrage avec fonctions non linéaires, ou bruits non gaussiens, l'IMM rencontre les mêmes difficultés que l'EKF et donne des résultats peu satisfaisants. S'ajoute à cela le problème difficile de "l'accrochage" du nouveau modèle à l'instant de manœuvre.

L'un des intérêts majeurs des méthodes particulières est justement de traiter indistinctement les problèmes quelles que soient les formes des bruits ou des fonctions (de dynamique ou d'observation). Les méthodes particulières sont

donc particulièrement adaptées aux cas multimodaux, dont le problème à dynamiques multiples fait partie. De plus, ces méthodes de filtrage ne fournissent pas simplement un estimateur particulier de l'état du mobile (espérance et covariance ou maximum de vraisemblance), mais elles approchent, de manière non paramétrique, la probabilité de présence du mobile en toute zone de l'espace d'état, sachant les mesures. Celle-ci constitue l'information maximale accessible et permet ensuite à l'utilisateur de choisir l'estimateur qui convient le mieux à son problème.

Le Multiple Model Particle Filter (MMPF) présenté, dans la suite, est une adaptation au problème à dynamiques multiples du Regularized Particle Filter (RPF) dont différentes versions sont décrites dans [9, 12, 8, 7]. Les performances du MMPF sont testées ici sur le problème de pistage par mesure d'azimut et de distance d'une cible manœuvrante.

2 Description du Filtre Optimal

On cherche donc à approcher à chaque instant la probabilité conditionnelle de l'état X_t sachant les mesures passées $(Y_1, \dots, Y_t) = Y_1^t$. On appelle filtre optimal, la suite des lois conditionnelles $(\mathcal{L}(X_t|Y_1^t))_{t \geq 1}$. Mais, s'ajoute au problème classique du filtrage, une difficulté supplémentaire puisque l'état du modèle, θ_t , est a priori inconnu. Pour palier ce problème, on introduit l'état augmenté $(X_t, \theta_t) = E_t$ (état du signal + état du modèle), dans le but d'estimer simultanément, à chaque instant, l'état de la cible et du modèle. La loi conditionnelle, $\mathcal{L}(X_t|Y_1^t)$, s'obtient alors comme une marginale de $\mathcal{L}(E_t|Y_1^t)$,

$$P(X_t|Y_1^t) = \sum_{k=1}^M P(X_t, \theta_t = k|Y_1^t) \quad (3)$$

L'évolution de $\mathcal{L}(E_t|Y_1^t)$ entre deux instants d'observation $t-1 \geq 1$ et t se déroule en deux étapes majeures.

A. Prédiction : $\mathcal{L}(E_{t-1}|Y_1^{t-1}) \rightarrow \mathcal{L}(E_t|Y_1^{t-1})$

$$P(E_t|Y_1^{t-1}) = \int_{e_{t-1}} P(E_t|E_{t-1} = e_{t-1})P(E_{t-1} \in de_{t-1}|Y_1^{t-1}) \quad (4)$$

L'originalité du cas à modèles multiples est que la transition, de E_{t-1} à E_t n'est pas simulable directement, elle requiert les simulations des transitions marginales, $\mathcal{L}(\theta_t|\theta_{t-1})$, puis $\mathcal{L}(X_t|X_{t-1}, \theta_t)$. En effet, comme θ_t est indépendant de X_{t-1} sachant θ_{t-1} ,

$$P(E_t|E_{t-1}) = P(X_t|\theta_t, X_{t-1})P(\theta_t|\theta_{t-1}) \quad (5)$$

B. Correction : $\mathcal{L}(E_t|Y_1^{t-1}) \rightarrow \mathcal{L}(E_t|Y_1^t)$

On déduit de la règle de Bayes et du fait que Y_t est indépendant de Y_1^{t-1} sachant l'état augmenté (X_t, θ_t) (cf. équation (2)),

$$P(E_t|Y_1^t) = \frac{P(Y_t|E_t)}{P(Y_t)}P(E_t|Y_1^{t-1}) \quad (6)$$

La loi conditionnelle, $\mathcal{L}(E_t|Y_1^t)$, s'obtient comme le produit normalisé de la loi prédite, par la vraisemblance de E_t par rapport à la mesure Y_t , $P(Y_t|E_t)$. Notons que, dans le cas du bruit gaussien, centré, de covariance R , on a, pour $e = (x, k)$,

$$P(Y_t|E_t = e) = \frac{1}{(2\pi)^{d/2}\sqrt{\det R}} \exp\left\{-\frac{1}{2}(Y_t - h_k(x)) * R^{-1}(Y_t - h_k(x))\right\}$$

3 Le Filtre Particulaire Classique

On le retrouve sous le nom de "Interacting Particle Filter", dans [3, 4], qui fournissent des analyses théoriques du comportement du filtre. Il est appelé aussi "Bootstrap Filter" dans [6] où le filtre est appliqué au problème du pistage par mesure d'angle. Par ailleurs, une version de ce filtre est développée pour l'étude du cas du temps continu dans [2].

Le filtre particulaire est essentiellement basé sur le principe de Monte Carlo. L'idée consiste à approcher la probabilité conditionnelle à l'aide d'un système de particules distribuées quasiment suivant cette loi. Pour cela, entre deux instants d'observation, on simule N trajectoires satisfaisant à l'évolution du filtre optimal, décrite ci-dessus. Supposons, qu'on dispose d'une mesure approchant la loi conditionnelle, $\mathcal{L}(E_{t-1}|Y_1^{t-1})$, et qu'on désire approcher $\mathcal{L}(E_t|Y_1^t)$. Le filtrage particulaire propose la démarche suivante,

A. Prediction échantillonnée (cf. équations (4,5))

1. Simuler N v.a. i.i.d., $(E_{t-1}^1, \dots, E_{t-1}^N)$, suivant l'approximation de $\mathcal{L}(E_{t-1}|Y_1^{t-1})$
2. Simuler, pour chaque v.a., E_{t-1}^i , une v.a., E_{t-1}^i , suivant $\mathcal{L}(E_t|E_{t-1} = E_{t-1}^i)$:
 $\rightarrow \theta_{t-1}^i \sim \mathcal{L}(\theta_t|\theta_{t-1} = \theta_{t-1}^i)$
 $\rightarrow X_{t-1}^i \sim \mathcal{L}(X_t|\theta_t = \theta_t^i, X_{t-1} = X_{t-1}^i)$

On obtient alors un nouveau système de particules prédites, $(E_{t-1}^1, \dots, E_{t-1}^N)$ qui nous permet d'approcher la loi prédite, $\mathcal{L}(E_t|Y_1^{t-1})$, par la loi discrète,

$$\frac{1}{N}\delta_{E_{t-1}^1} + \dots + \frac{1}{N}\delta_{E_{t-1}^N} \quad (7)$$

B. Correction (cf. équation (6))

1. Pondérer chaque élément, E_{t-1}^i , du système de particules prédites, par un poids proportionnel à sa vraisemblance (fonction de l'observation courante Y_t),

$$p_t^i = \frac{P(Y_t|E_t = E_{t-1}^i)}{\sum_{j=1}^N P(Y_t|E_t = E_{t-1}^j)} \quad (8)$$

On approche alors la loi conditionnelle, $\mathcal{L}(E_t|Y_1^t)$, par la loi discrète pondérée,

$$p_t^1\delta_{E_{t-1}^1} + \dots + p_t^N\delta_{E_{t-1}^N} \quad (9)$$

A l'instant suivant, $t+1$, dans la phase (A. 1.), on simule un échantillon issu de l'approximation particulaire de $\mathcal{L}(E_t|Y_1^t)$ donnée à l'instant, t , par (9). Cette phase, appelée redistribution, a pour effet de multiplier les particules les plus vraisemblables par rapport à l'observation courante et d'éliminer les autres, en gardant un nombre constant de particules. Ainsi, le système se concentre naturellement dans les zones intéressantes de l'espace d'état.

L'inconvénient de la redistribution est de produire un nouveau système, où plusieurs particules sont confondues. Si le bruit de dynamique est faible ou inexistant, les redistributions répétées appauvrissent de plus en plus le système de particules, jusqu'à le réduire à un seul point. Ce phénomène de "dégénérescence des particules" limite voire élimine totalement la capacité d'exploration du filtre.

De plus, à l'étape de correction, le caractère discret de l'approximation particulaire limite la mise à jour aux seuls points du système de particules. Ainsi, si la fonction de vraisemblance concentre sa masse dans une zone de l'espace d'état extérieure au système de particules (cas du faible bruit de mesure, où la zone de forte vraisemblance est trop réduite pour rencontrer le système de particules), la correction néglige cette zone (la plus intéressante), et le filtre particulaire diverge.

4 Le Filtre Particulaire Régularisé

On a mis en évidence, lors d'applications au pistage, ([9]), les comportements critiques du filtre particulaire classique annoncés ci-dessus. Pour palier ces problèmes, on a proposé les Filtres Particulaires Régularisés, (RPF), dans lesquels on rajoute une étape, dite de régularisation, à l'algorithme particulaire classique. Il existe, actuellement, plusieurs versions de RPF. On peut les ranger dans deux grandes catégories (pre-RPF et post-RPF) suivant qu'on insère l'étape de régularisation avant ou après l'étape de correction. Pour une analyse théorique de ces méthodes, on peut se reporter à [8]. Dans [9, 12], on décrit une version de filtre régularisé après correction, que l'on compare à l'EKF notamment pour le problème de pistage par mesure

d'angle seul. Par ailleurs, Hürzeler et Künsch ont proposé, dans [7], une version de filtre régularisé avant correction, le Kernel Filter. Ce dernier est adapté, pour le temps réel, dans [10, 11], puis appliqué au problème délicat du pistage haute précision.

4.1 Principe de Régularisation

On appelle noyau de régularisation, un opérateur, $K : \mathbb{R}^d \rightarrow \mathbb{R}^+$ tel que :

$$K \geq 0, \int K(x)dx = 1, \int xK(x)dx = 0$$

On appelle facteur de régularisation, un réel $h > 0$, qui caractérise le noyau dilaté,

$$K_h(x) = \frac{1}{h^d} K\left(\frac{x}{h}\right) \quad \forall x \in \mathbb{R}^d$$

Si on note $*$ le produit de convolution, et ν une mesure de probabilité, la régularisée de ν est la mesure de probabilité absolument continue,

$$\{\nu * K_h\} = \int_{u \in \mathbb{R}^d} \nu(du) K_h(\cdot - u)$$

Si ν est une mesure discrète,

$$\nu = p^1 \delta_{X^1} + \dots + p^N \delta_{X^N}$$

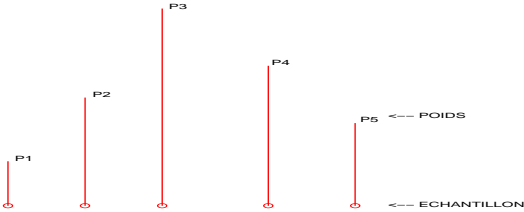


FIG. 1: Mesure discrète pondérée

La densité de la mesure régularisée, $\{\nu * K_h\}$, est simplement la somme des noyaux dilatés, pondérés et translatés,

$$p^1 K_h(\cdot - X^1) + \dots + p^N K_h(\cdot - X^N)$$

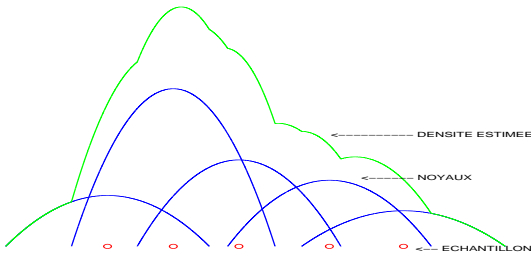


FIG. 2: Densité de la régularisée

4.2 Régularisation et Filtrage Particulaire

Lorsqu'on régularise après correction (post-RPF), l'algorithme est constitué de ces trois étapes successives, A. Prédiction, B. Correction, C. Régularisation.

On approche $\mathcal{L}(X_t|Y_1^t)$ non plus par la mesure discrète, marginale de (9),

$$\{p_t^1 \delta_{X_{t|t-1}^1} + \dots + p_t^N \delta_{X_{t|t-1}^N}\}$$

mais par la mesure de densité,

$$\{p_t^1 K_h(\cdot - X_{t|t-1}^1) + \dots + p_t^N K_h(\cdot - X_{t|t-1}^N)\}$$

La phase de redistribution est alors remplacée par un échantillonnage suivant une loi continue. Ainsi, le nouveau système de particules est assuré de contenir N particules différentes, ce qui règle le problème de "dégénérescence des particules".

Lorsqu'on régularise avant correction (pre-RPF), l'algorithme est constitué de ces trois étapes successives, A. Prédiction, B. Régularisation, C. Correction.

On approche la mesure prédite $\mathcal{L}(X_t|Y_1^{t-1})$ non plus par la mesure discrète, marginale de (7),

$$\frac{1}{N} \{\delta_{X_{t|t-1}^1} + \dots + \delta_{X_{t|t-1}^N}\}$$

mais par la mesure de densité,

$$\left\{ \frac{1}{N} K_h(\cdot - X_{t|t-1}^1) + \dots + \frac{1}{N} K_h(\cdot - X_{t|t-1}^N) \right\}$$

A l'étape de correction, on met à jour l'infinité de points du support de la densité régularisée, alors qu'en filtrage classique on se contente de pondérer les N particules du système. Ainsi, la correction ne néglige plus de zones de l'espace d'état.

4.3 Choix Optimal des Paramètres de Régularisation

Pour minimiser la distance entre la vraie loi, $\mathcal{L}(X_t|Y_1^t)$, et l'approximation régularisée, on utilise la théorie de l'estimation de densité ([5, 13]) qui nous permet d'optimiser, sous certaines contraintes de régularité du noyau, K , la forme du noyau, ainsi que la valeur du facteur, h . On approche les vrais paramètres optimaux, par des paramètres optimaux empiriques dépendants de l'échantillon, $(X_{t|t-1}^1, \dots, X_{t|t-1}^N)$,

$$K_{opt}(x) = \begin{cases} \frac{d+2}{2c_d} (1 - x^t x) & \text{si } x^t x < 1 \\ 0 & \text{sinon} \end{cases}$$

$$h_{opt} = \left[\frac{4d}{N} \frac{\sqrt{\det(\Gamma)}}{\text{trace}(\Gamma^{-1})^2 + 2\|\Gamma^{-1}\|^2} \right]^{\frac{1}{d+4}}$$

Γ désigne la matrice de covariance de l'échantillon, (X^1, \dots, X^N) , et c_d le volume de la sphère unité de \mathbb{R}^d . Le K_{opt} obtenu, ici est communément appelé le "noyau d'Epanechnikov". Dans la pratique, pour des raisons de robustesse (gestion des cas multimodaux), on choisit $h = h_{opt}/2$. Pour alléger le coût des simulations suivant la mesure régularisée, on peut toujours choisir un noyau gaussien.

5 Applications

Les performances de ce nouveau filtre ont été testées sur le problème de pistage par mesure d'azimut et de distance d'une cible manœuvrante. Les résultats rapportés ci-dessous sont moyennés sur 100 simulations réalisées avec 5000 particules. L'observateur est fixe au point (0,0). La cible est située, initialement, au point (50km,50km). Elle se déplace, dans le plan, selon l'axe (oy), à vitesse constante (5m/s), pendant 250s, ensuite elle effectue un virage à la vitesse constante, $\Omega=0.006\text{rad/s}$, pendant 500s, pour reprendre enfin un mouvement rectiligne uniforme.

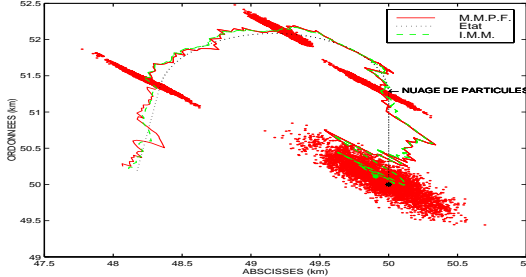


FIG. 3: Trajectoires réelles et estimées

On mesure toutes les 5s l'azimut et la distance à l'observateur, avec des bruits gaussiens d'écart type = 0.5° et 20m. La matrice de transition de la chaîne de Markov des modèles ($\theta=1$: modèle rectiligne uniforme, et $\theta=2$: modèle de rotation) est la suivante,

$$P(\theta_0 = 1) = 0.98, \quad [p(i, j)] = \begin{bmatrix} 0.98 & 0.02 \\ 0.02 & 0.98 \end{bmatrix}$$

On compare les performances du MMPF à celles d'un IMM. On donne à l'IMM la vraie vitesse de rotation, Ω alors que le MMPF l'ignore et l'initialise la vitesse du virage uniformément dans l'intervalle, $[-2\Omega, 2\Omega]$.

On observe, ci-dessous, un très bon comportement du MMPF, pour un temps de calcul 7 fois supérieur à l'IMM (localisation rapide de l'instant de manœuvre, très bonne estimation de la vitesse de rotation lors du virage).

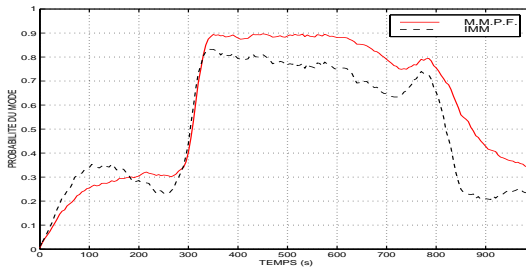


FIG. 4: Probabilités du virage estimées

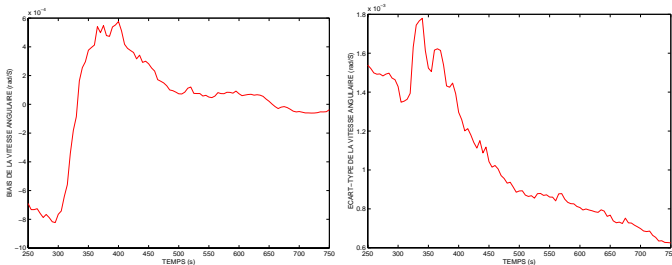


FIG. 5: Biais et Std de la vitesse angulaire estimée

Références

- [1] A. P. Henk Blom and Yaakov Bar-Shalom. The interacting multiple model algorithm for systems with markovian switching coefficients. *IEEE Transactions On Automatic Control*, (8):780–783, 1988.
- [2] D. Crisan and T. Lyons. Nonlinear filtering and measure valued processes. *Probability Theory and Related Fields*, (109):217–244, 1997.
- [3] P. Del Moral. Nonlinear filtering: Interacting particle resolution. *Probability Theory/Statistics*, 325:653–658, 1997.
- [4] P. Del Moral and A. Guionnet. Large deviation for interacting particle systems. application to nonlinear filtering problems. In Université Paul Sabatier, editor, *Publication du Laboratoire de Statistique et Probabilités*, volume 05-97. Toulouse, 1997.
- [5] L. Devroye. *A Course on Density Estimation*, volume 14 of *Progress in Probability and Statistics*. Birkhäuser, 1987.
- [6] N. Gordon, D. Salmond, and A. Smith. Novel approach to nonlinear/non-gaussian bayesian state estimation. In *IEE Proceedings*, volume 140, pages 107–113, April 1993.
- [7] M. Hürzeler and H. R. Künsch. Monte carlo approximations for general state-space models. *Journal of Computational and Graphical Statistics*, (7):175–193, 1998.
- [8] F. Le Gland, C. Musso, and N. Oudjane. An analysis of regularized interacting particle methods for nonlinear filtering. In *IEEE European Workshop on Computer-Intensive Methods in Control and Signal Processing, Prague, Proceedings*, September 1998.
- [9] C. Musso and N. Oudjane. Regularisation schemes for branching particle systems as a numerical solving method of the nonlinear filtering problem. In *Irish Signals Systems Conference, Dublin, Proceedings*, June 1998.
- [10] C. Musso and N. Oudjane. Méthodes statistiques en filtrage particulaire appliqué au pistage. In *International Conference on Radar Systems, Brest, France, Proceedings, (à paraître)*, May 1999.
- [11] C. Musso and N. Oudjane. Particle methods for multimodal filtering. In *International Conference on Information Fusion, Silicon Valley, CA, USA, Proceedings*, July 1999.
- [12] N. Oudjane and C. Musso. Regularized particle schemes applied to the tracking problem. In *International Radar Symposium, Munich, Proceedings*, September 1998.
- [13] B. W. Silverman. *Density Estimation for Statistics and Data Analysis*. Number 26 in Monographs on Statistics and Applied Probability. Chapman and Hall, 1986.