Traffic intensity estimation via PHD filtering

G. Battistelli [1], L. Chisci [1], S. Morrocchi [1], F. Papi [1], A. Benavoli [2], A. Farina [2], A. Graziano [2]

[1] *DSI, Università di Firenze, Florence, Italy*
e-mail: {*battistelli,chisci*}*@dsi.unifi.it*

[2] *Engineering Division, SELEX Sistemi Integrati, Rome, Italy*
e-mail: {*abenavoli,afarina,agraziano*}*@selex-si.com*

***Abstract*— The paper will address the estimation of road traffic intensity from available measurements of mobile vehicles' coordinates. To this end, the work will jointly exploit PHD (Probability Hypothesis Density) filtering techniques based on the so called particle filter approach and road-map information.**

I. INTRODUCTION

Traffic intensity estimation, in both urban and extra-urban scenarios, is important for on-line road traffic management. Nowadays, wireless communication technologies may provide measurements of vehicles' positions at different levels of accuracy as well as with different probabilities of vehicle detection and false alarm []. This makes possible to exploit the well established multitarget tracking techniques developed by the radar data processing [1] and information fusion community for road traffic monitoring purposes. In particular, the recent work on the *Probability Hypothesis Density* (PHD) filter seems especially well suited to crowded scenarios like the one arising from road traffic. A further motivation for resorting to PHD filtering techniques is that the objective of traffic monitoring is not estimating the state (position and velocity) of any individual vehicle but rather estimating the spatial distribution (i.e. traffic intensity or density) of the overall vehicle population. Based on the above motivations, the contribution of the present paper is twofold: on one hand to show how the knowledge of the road-map can be exploited in the development of the vehicle motion, birth and death models of the PHD filter and, on the other hand, to evaluate how the performance of the PHD-based traffic intensity estimation depends on the characteristics of the available measurements (i.e. position accuracy, detection and false alarm probabilities).

II. PHD FILTERING

In multitarget and/or multisensor tracking there are essentially two approaches to cope with the fact that the source (either a target or clutter) originating a given measurement is unknown. The traditional approach [3] consists of tracking each individual target with a separate filter and, thus, requires explicit association of the available measurements to the detected targets. Due to its combinatorial nature, association typically represents the most computationally intensive task of a tracking system and may provide poor performance for large numbers of closely-spaced measurements and/or targets. An alternative approach is that of regarding targets and measurements as *random sets* [4]-[5], i.e. objects in which randomness is not only in the assumed values but also in the number of elements. In this framework, the objective is to recursively estimate the random target set $\mathbf{X}_t = \{\mathbf{x}_{t,1}, \mathbf{x}_{t,2}, \ldots, \mathbf{x}_{t,n_t}\}$, i.e. the set of the states of all targets that are present at time t, given the random measurements sets $\mathbf{Y}_k = \{\mathbf{y}_{k,1}, \mathbf{y}_{k,2}, \ldots, \mathbf{y}_{k,m_k}\}$, i.e. the sets of all measurements collected from all sensors at time k, for all k up to time t. In [6] the multi-target Bayes filter recursion propagating the probability density function $p\left(\mathbf{X}_t | \mathbf{Y}^t\right)$, where $\mathbf{Y}^t \triangleq \{\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_t\}$, has been provided by exploiting random set statistics [5]. This approach, though theoretically optimal, involves integrations in infinite dimensional spaces and is, therefore, intractable in most practical applications. To make the multi-target Bayes filter computationally feasible, the so called *Probability Hypothesis Density (PHD)* filtering approach has been proposed [6]. The idea underlying the PHD filter is to propagate a suitable *density function* $D(\mathbf{x})$ in the target state-space $\mathcal{X} \subset \mathbb{R}^n$ ($n = dim\ \mathbf{x}$) such that, for any region $\mathcal{S} \subseteq \mathcal{X}$, the expected number of targets in $\mathcal{S}$ is given by

$$n(\mathcal{S}) \; = \; \int_{\mathcal{S}} D\left(\mathbf{x}\right) \, d\mathbf{x}, \qquad \text{(II.1)}$$

i.e. by integration of $D(\cdot)$ over $\mathcal{S}$. Hereafter, $D(\cdot)$ will be referred to as PHD function and the objective of the PHD filter is clearly the time propagation

$$D_{t-1|t-1}(\mathbf{x}) \; \rightarrow \; D_{t|t-1}(\mathbf{x}) \; \rightarrow \; D_{t|t}(\mathbf{x})$$

where $D_{t|t-1}(\cdot)$ and $D_{t|t}(\cdot)$ denote the PHDs at time t based on $\mathbf{Y}^{t-1}$ and, respectively, $\mathbf{Y}^t$. In order to provide the PHD recursion, let us introduce the following notation:

- $p_{t|t-1}(\mathbf{x}|\boldsymbol{\xi})$: single-target state transition PDF originated by the target dynamics $\mathbf{x}_{t+1} = \mathbf{f}(\mathbf{x}_t) + \mathbf{w}_t$;
- $\ell_t(\mathbf{x}|\mathbf{y})$: likelihood function originated by the measurement relationship $\mathbf{y}_t = \mathbf{h}(\mathbf{x}_t) + \mathbf{v}_t$;
- $P_{s,t}(\mathbf{x})$: survival probability of a target at time t and state $\mathbf{x}$;
- $P_{d,t}(\mathbf{x})$: detection probability of a target at time t and state $\mathbf{x}$;
- $b_t(\mathbf{x})$: *birth* density in the single-target state space;
- $c_t(\mathbf{y})$: *clutter* (false alarm) density in the single-measurement space $\mathcal{Y} \subset \mathbb{R}^p$ ($p = dim\ \mathbf{y}$).

Then, under reasonable assumptions [6] on the target dynamics as well as on the measurements generation, the PHD recursion takes the following form:

$$\begin{aligned} D_{t|t-1}(\mathbf{x}) &= b_t(\mathbf{x}) + P_{s,t}(\mathbf{x}) \int p_{t|t-1}(\mathbf{x}|\boldsymbol{\xi}) D_{t|t-1}(\boldsymbol{\xi}) d\boldsymbol{\xi} \\ D_{t|t}(\mathbf{x}) &= \left[1 - P_{d,t}(\mathbf{x})\right] D_{t|t-1}(\mathbf{x}) + \\ &+ \sum_{\mathbf{y} \in \mathbf{Y}_t} \frac{P_{d,t}(\mathbf{x}) \ell_t(\mathbf{x}|\mathbf{y}) D_{t|t-1}(\mathbf{x})}{c_t(\mathbf{y}) + \int P_{d,t}(\boldsymbol{\xi}) \ell(\boldsymbol{\xi}|\mathbf{y}) D_{t|t-1}(\boldsymbol{\xi}) d\boldsymbol{\xi}} \end{aligned} \tag{II.2}$$

Notice that, with respect to the general PHD recursion derived in [6], the target *spawning* has not been included in (II.2) as this phenomenon is negligible in the road traffic scenario considered in this work. It is also worth pointing out that the PHD recursion (II.2) does not admit, in general, a closed-form solution just like the nonlinear and/or non-Gaussian single-target Bayes filter. The current state-of-art provides two approaches to the implementation of the PHD filter, i.e. the *Gaussian Mixture* PHD (GM-PHD) filter [7] and the *Particle Filter* PHD (PF-PHD) filter [8]. While the GM-PHD filter provides an analytical solution of the PHD recursion under suitable restrictive assumptions, the PF-PHD filter yields only a numerical approximation (via sequential Monte Carlo integration methods) of such a recursion but has general applicability. In this work the PF implementation of PHD will be adopted.

III. The traffic scenario

A. Road network

In a real scenario, a vehicle can move along a complicated network of roads intersecting at junctions. Hence, it is important to suitably represent the *road network* as well as to handle in a smart way the multiple hypotheses on the target's behavior in the proximity of the road junctions.

To this end, a real curved road can be approximated, to any degree of accuracy, by means of a *polyline* consisting of multiple, say N, line segments. Hence, a road (polyline) is specified by the sequence of endpoints $(x^0, y^0), (x^1, y^1), \ldots, (x^N, y^N)$ of the road (line) segments. Segment i, joining (x^{i-1}, y^{i-1}) and (x^i, y^i), can therefore be described by the line equation $\alpha^i x + \beta^i y = \gamma^i$ with coefficients

$$\alpha^i = y^i - y^{i-1}, \beta^i = x^{i-1} - x^i, \gamma^i = y^i x^{i-1} - x^i y^{i-1}.$$

Along these lines, the topology of a road network can be described as a directed graph whose nodes and arcs represent the *junctions* and, respectively, the *roads*. Let $\mathcal{J}$ and $\mathcal{R}$ denote the sets of junctions (nodes) and, respectively, roads (arcs). For each road $r \in \mathcal{R}$ the following information must be given: the junctions $j_1(r)$ and $j_2(r)$ delimiting the road; the number $N(r)$ of line segments making up the road; the endpoints $\left\{ \left(x^i(r), y^i(r)\right) \right\}_{i=0}^{N(r)}$ from which the line coefficients $\left\{ \left(\alpha^i(r), \beta^i(r), \gamma^i(r)\right) \right\}_{i=1}^{N(r)}$ can easily be recovered. Note that in the present model the traffic on road r is supposed to flow from junction $j_1(r)$ to junction $j_2(r)$. Then 2-way roads are modelled as two arcs of the graph with opposite directions. For each junction $j \in \mathcal{J}$, one needs to specify: the set $\mathcal{R}_-$ of roads going into the junction; the set $\mathcal{R}_+(j)$ of roads coming out of the junction; the junction's position $(x(j), y(j))$.

B. Vehicle motion and birth-death modelling

The motion of a vehicle along a road segment can be modelled by means of a 2-D *continuous white noise acceleration (CWNA)* kinematic model with 4-dimensional state vector $\mathbf{x} = [x, \dot{x}, y, \dot{y}]'$ where x, y and $\dot{x}, \dot{y}$ are the cartesian position and, respectively, velocity components. The resulting state equation for a segment with line coefficients (α, β, γ) turns out to be

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{w}_t$$

where

$$\mathbf{A} = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$\mathbf{w}_t$ is a zero mean white-noise with variance

$$\mathbf{Q} = \frac{\sigma_s^2}{\alpha^2 + \beta^2} \begin{bmatrix} \frac{\beta^2 T^3}{3} & \frac{\beta^2 T^2}{2} & -\frac{\alpha\beta T^3}{3} & -\frac{\alpha\beta T^2}{2} \\ \frac{\beta^2 T^2}{2} & \beta^2 T & -\frac{\alpha\beta T^2}{2} & -\alpha\beta T \\ -\frac{\alpha\beta T^3}{3} & -\frac{\alpha\beta T^2}{2} & \frac{\alpha^2 T^3}{3} & \frac{\alpha^2 T^2}{2} \\ -\frac{\alpha\beta T^2}{2} & -\alpha\beta T & \frac{\alpha^2 T^2}{2} & \alpha^3 T \end{bmatrix},$$

T is the measurement sampling interval and σ_s is the standard deviation of the target's speed. Notice that such a motion model imposes the fulfillment of the linear constraint $\alpha x + \beta y = \gamma$ for any possible noise realization.

As far as road segment update is concerned, let us assume that the measurement sampling interval is fast enough with respect to the lengths of the road segments so that the transition can only occur between contiguous segments. Let r_t and i_t denote the road and, respectively, the specific segment at time t. Given the position (x_{t+1}, y_{t+1}) the road segment update can be carried out as follows

$$i_{t+1} = \begin{cases} i_t, & 0 \le \lambda_{t+1} \le 1 \\ i_t + 1, & \lambda_{t+1} > 1 \text{ and } i_t < N(r_t) \\ 1, & \lambda_{t+1} > 1 \text{ and } i_t = N(r_t) \end{cases}$$

where

$$\lambda_{t+1} = \frac{x_{t+1} - x^{i_t - 1}(r_t)}{x^{i_t}(r_t) - x^{i_t - 1}(r_t)}.$$

When $\lambda_{t+1} > 1$ and $i_t = N(r_t)$ (i.e., when the vehicle enters junction $j_2(r_t)$) also the road index r_t needs to be updated. Specifically, after reaching junction j, a vehicle may either enter a road belonging to $\mathcal{R}_+(j)$ or exit the road network. This can be modelled by defining suitable transition probabilities $\pi_{r,k}(j)$ for each pair of roads k and r entering and, respectively, exiting junction j. Then, the probability that a vehicle leaves the road network at junction j when coming from road r turns out to be $1 - \sum_{k \in \mathcal{R}_+(j)} \pi_{r,k}(j)$.

After the new road r_{t+1} and the new road segment i_{t+1} have been determined, the state vector $\mathbf{x}_{t+1}$ is updated so as to ensure the fulfilment of the corresponding equality constraint.

The description of the road scenario is completed by specifying the road network inflow model (i.e., the birth density). To this end, let $\mathcal{A} \subseteq \mathcal{J}$ be the set of access points to the road network. For each road r departing from an access point j, $B_r(j)$ denotes the expected number of vehicles entering the road network at road r in each sampling interval. It is supposed that the initial state of a vehicle entering the road network at road r is distributed according to some probability density function $p_r(\mathbf{x})$ (e.g., a truncated Gaussian distribution can be used).

C. Measurements characteristics

A number of mobile-based location systems [] are currently available that provide measurements of mobile coordinates with respect to a set of base stations with known positions. Such systems are usually classified according to the number of base stations with which the mobile can communicate, the most common situations being single-site visibility or three-site visibility. In this paper the former case is considered, wherein Angle of Arrival (AOA) and Time of Arrival (TOA) are determined at a single base station, thus providing measurements of range ρ_t and azimuth θ_t. Then the measurement equation is given by

$$\mathbf{y}_t = \mathbf{h}(\mathbf{x}_t) + \mathbf{v}_t \qquad \text{(III.1)}$$

where $\mathbf{y}_t = [\rho_t,\ \theta_t]'$ is the measurement vector at time t, the nonlinear function $\mathbf{h}(\cdot)$ is defined as

$$\mathbf{h}(\mathbf{x}) = \begin{bmatrix} \sqrt{x^2+y^2} \\ \angle\,(x+iy) \end{bmatrix},$$

and $\mathbf{v}_t$ is a zero-mean white Gaussian measurement noise with covariance matrix $\mathbf{R} = \mathrm{diag}\{\sigma_\rho, \sigma_\theta\}$.

For the sake of simplicity, it is supposed that each vehicle can be detected at each time instant t with a certain probability P_d independently of its position. Further, it is supposed that, at each time instant t, on average n_c clutter measurements are generated with an uniform distribution in the measurement space $\mathcal{Y}$.

IV. PF-PHD ALGORITHM FOR TRAFFIC INTENSITY ESTIMATION

In this section a Particle Filter algorithm is proposed for the approximate propagation of the PHD in the considered road traffic scenario. To this end, at each time instant t the PHD $D_{t|t}(\cdot)$ is approximated by a set of particles $\left\{(w_t^k, \mathbf{x}_t^k)\right\}_{k=1}^{N_{t|t}}$ where $\mathbf{x}_t^k \in \mathcal{X}$ and $w_t^k \in \mathbb{R}$ represent the state and, respectively, the weight of the k-th particle. The approximation is designed so that the cumulative weight of particles in a region $\mathcal{S}$ of the state space represents the expected number of targets in such a region, i.e.,

$$\sum_{\mathbf{x}_t^k \in \mathcal{S}} w_t^k \approx \int_{\mathcal{S}} D_{t|t}(\mathbf{x})\ d\mathbf{x}\,.$$

Then the estimated number of vehicles on the road network at time t can be obtained as

$$\hat{n}_{t|t} = \sum_{k=1}^{N_{t|t}} w_t^k\,.$$

The PF-PHD algorithm consists mainly of three steps: i) a prediction step wherein the particles $\left\{(w_{t-1}^k, \mathbf{x}_{t-1}^k)\right\}_{k=0}^{N_{t-1|t-1}}$ representing $D_{t-1|t-1}(\cdot)$ are propagated according to the vehicle motion model and new particles are generated on the basis of the birth model; ii) an update step wherein the particle weights are modified according to the likelihood with respect to the measurements set $\mathbf{Y}_t$; iii) a resampling step [2] wherein $N_{t|t} = N_p\,\hat{n}_{t|t}$ particles are sampled from the existing ones to form a new set $\left\{(w_t^k, \mathbf{x}_t^k)\right\}_{k=1}^{N_{t|t}}$ representing $D_{t|t}(\cdot)$ so that on average N_p particles are allocated to each vehicle. Specifically, at each time instant the following algorithm is applied.

PF-PHD algorithm

step 1 (prediction):
for $k = 1, \ldots, N_{t-1|t-1}$
  sample $\mathbf{x}_t^k$ from $\mathbf{x}_{t-1}^k$ according to the vehicle motion model defined in Section III;
  set $w_t^k = 0$ if the particle exits the road network or $w_t^k = w_{t-1}^k$ otherwise;
end
for $j \in \mathcal{A}$ and $r \in \mathcal{R}_+(j)$
  generate $B_r(j)N_p$ new particles according to the PDF $p_r(\cdot)$ with weights $1/N_p$;
end
set $N_{t|t-1} = N_{t-1|t-1} + \sum_{j\in\mathcal{A}} \sum_{r\in\mathcal{R}_+(j)} B_r(j)N_p$;
set $\hat{n}_{t|t-1} = \sum_{k=1}^{N_{t|t-1}} w_t^k$;
step 2 (update):
for $l = 1, \ldots, m_t$
  for $k = 1, \ldots, N_{t|t-1}$
    compute the likelihood $\lambda_l^k = \ell(\mathbf{x}_t^k | \mathbf{y}_{t,l})$;
  end
  compute the normalization factor $\Lambda_l = \sum_{k=1}^{N_{t|t-1}} \lambda_l^k$;
end
for $k = 1, \ldots, N_{t|t-1}$
  set $w_t^k = w_t^k \left(1 - P_d + \frac{m_t - n_c}{m_t} \sum_{l=1}^{m_k} \lambda_l^k / \Lambda_l\right)$;
end
step 3 (resampling):
set $\hat{n}_{t|t} = \sum_{k=1}^{N_{t|t-1}} w_t^k$;
set $N_{t|t} = N_p\,\hat{n}_{t|t}$;
resample $N_{t|t}$ particles from $\left\{(w_t^k, \mathbf{x}_t^k)\right\}_{k=1}^{N_{t|t-1}}$ to get $\left\{(w_t^k, \mathbf{x}_t^k)\right\}_{k=1}^{N_{t|t}}$;

It is immediate to check that the estimated number of vehicles $\hat{n}_{t|t}$ turns out to be

$$\hat{n}_{t|t} = (1 - P_d)\hat{n}_{t|t-1} + m_t - n_c\,. \qquad \text{(IV.1)}$$

This is due to the presence of the scale factor $(m_t - n_c)/m_t$ in the computation of the particle weights in the update step. Such a scale factor represents the expected fraction of non-clutter measurements at time t and is needed to ensure that $\hat{n}_{t|t}$ is an unbiased estimate of the number of vehicles.

As well known, the resampling step is needed to avoid an uncontrolled growth in the number of particles. This is usually achieved by eliminating particles with low weights and multiplying particles with high weights. Unfortunately, this may lead to some problems when the detection probability is low since consecutive missed detections for an isolated vehicle may lead to the disappearance of all the particles in its proximity. To mitigate such a drawback, at each time instant a certain number of particles are generated from the measurements that are close to the road network (i.e., probably not clutter) but far from all the particles (note that this can be seen as some sort of importance sampling). Specifically, let $d(\mathbf{y},\mathcal{R})$ denote the distance of a measurement $\mathbf{y} \in \mathcal{Y}$ from the road network. Then, the update step of the PF-PHD algorithm is modified as follows.

---

step 2 (update):
for $l = 1, \ldots, m_t$
  for $k = 1, \ldots, N_{t|t-1}$
    compute the likelihood $\lambda_l^k = \ell(\mathbf{x}_t^k|\mathbf{y}_{t,l})$;
  end
  compute the normalization factor $\Lambda_l = \sum_{k=1}^{N_{t|t-1}} \lambda_l^k$;
  set $\mu_l = \varphi(\Lambda_l)$ if $d(\mathbf{y}_{t,l},\mathcal{R}) < \varepsilon$ or μ_l = 0 otherwise;
end
for $k = 1, \ldots, N_{t|t-1}$
  set $w_t^k = w_t^k \left(1 - P_d + \frac{m_t - n_c}{m_t} \sum_{l=1}^{m_k} \lambda_l^k \varphi(\Lambda_l)/\Lambda_l\right)$;
end
for $l = 1, \ldots, m_t$
  if $\mu_l > 0$ generate N_p new particles on the road segment nearest to $\mathbf{y}_{t,l}$ with weights $(1 - \varphi(\Lambda_l))/N_p$;
  if $\mu_l > 0$ set $N_{t|t-1} = N_{t|t-1} + N_p$;
end

The function $\varphi(\cdot)$ is monotonically non-decreasing and takes value in the interval $[0, 1]$. Thus, the farther the measurement $\mathbf{y}_{t,l}$ is from the existing particles (i.e., the smaller is the total likelihood Λ_l), the greater is the weight $(1 - \varphi(\Lambda_l))/N_p$ assigned to the new-born particles. Finally, note that the weights of the existing particles are updated so that the cumulative sum of the particle weights (i.e., the expected number of vehicles) can still be obtained as in (IV.1).

V. Performance evaluation

A simple road scenario has been considered (see fig. 3). The following parameters have been chosen: $T = 5s$, $P_d = 0.2$, $n_c = 10$, $\sigma_\rho = 50\ m$, $\sigma_\theta = 1°$, $\sigma_s = 3\ m/s$, $N_p = 20$. Notice that, in steady state, the scenario has about 50 vehicles. The following performance metrics, averaged over 25 independent Monte Carlo runs, have been evaluated: the *estimation error* on the number of vehicles $n_{t|t} - n_t$ (see fig. 1) and the *dispersion index* (see fig. 2)

$$DI = \frac{1}{n_t} \frac{1}{N_p} \sum_{i=1}^{n_t} \sum_{k \in \mathcal{K}_t(i)} \delta\left(\mathbf{x}_{t,i}, \mathbf{x}_t^k\right)$$

where $\delta(\cdot,\cdot)$ is the Euclidean distance and $\mathcal{K}_t(i)$ is the set of the N_p particles closest to the ith vehicle at time t. Fig. 3 provides a 3D plot of the PHD function at a given time instant. The performance's dependence on parameters $P_d, n_c, \sigma_\rho, \sigma_\theta$ will be presented in the final version of the paper.

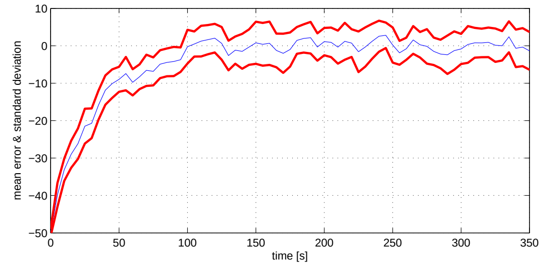

Fig. 1. Vehicle number mean estimation error and $\pm$ std. dev. band

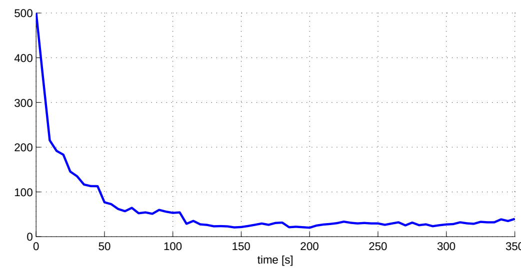

Fig. 2. Time behaviour of the dispersion index

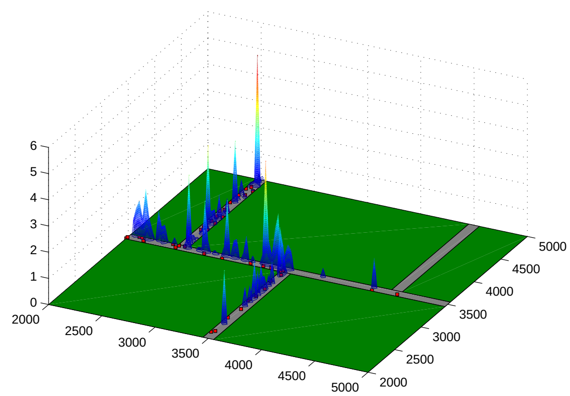

Fig. 3. 3D plot of the PHD function

References

[1] A. Farina, F.A. Studer, *Radar Data Processing, vol. I: Introduction and Tracking.* John Wiley & Sons, New York, 1985.
[2] B. Ristic, S. Arulampalam, N. Gordon, *Beyond the Kalman filter: particle filters for tracking applications*, Artech House, Norwood, MA, 2004.
[3] Y. Bar-Shalom and X.R. Li: *Multitarget-multisensor tracking: principles and techniques*, YBS Publishing, Storrs, CT, 1995.
[4] H.T. Nguyen: *An introduction to random sets*, Chapman & Hall, 2006.
[5] R.P.S. Mahler: *Statistical multisource multitarget information fusion*, Artech House, 2007.
[6] R.P.S. Mahler: "Multitarget Bayes filtering via first-order multitarget moments", *IEEE Trans. on Aerospace and Electronic Systems*, vol. 39, n. 4, pp. 1152-1178, 2003.

[7] B.N. Vo and W.K. Ma: “The Gaussian mixture Probability Hypothesis Density filter”, *IEEE Trans. on Signal Processing*, vol. 54, n. 11, pp. 4091-4104, 2006.
[8] B.N. Vo, S. Singh and A. Doucet: “Sequential Monte Carlo methods for multi-target filtering with random finite sets”, *IEEE Trans. on Aerospace and Electronic Systems*, vol. 41, n. 4, pp. 1224-1245, 2005.