Multi object tracking with PHD filter

Guillaume POIDATZ Runkun LUO

I. Introduction

A. Why PHD filter ?

If the number of targets is known and constant, there exists other well-known techniques that use one kalman filter (KF) per track and associate measurements to tracks such as global nearest neighbour (GNN). But these techniques are inconsistent with unkown or time-varying number of tracks. Moreover, they don't give precision on the confidence into the existence of the targets. This is the motivation for a new algorithm : Probability Hypothesis Density (PHD) based on random finite sets.

B. Random finite set

A random finite set is defined as a set containing a random number of random variables :

$$X = x_1, ..., x_n \tag{1}$$

where both the cardinality n and the x_i are randomly distributed

The cardinality distribution - that is the probability that X contains exactly i elements - is denoted as ρ and defined as follow :

$$\rho(i) = \frac{1}{i!} p(x_1, ..., x_i) \tag{2}$$

The probability density function (PDF) of X is then :

$$\pi(X) = \sum_{i=0}^{\infty} \rho(i) \tag{3}$$

C. predictive model

The predictive model defines the random finite set at time k knowing k-1 ($X_{k|k-1}$).

$$X_{k|k-1} = S_{k|k-1} \cup B_{k|k-1} \tag{4}$$

With S the survivors, and B the birth model

Considering that targets are independent with $X_k = \{x_{k-1}^1, ..., x_{k-1}^{n_{k-1}}\}$ the equation becomes :

$$X_{k|k-1} = B_{k|k-1} \cup S_k(x_{k-1}^1) \cup ... \cup S_k(x_{k-1}^{n_{k-1}}) \tag{5}$$

And the PDF of $X_{k|k-1}$ is then :

$$\pi_{X_k}(X_{k|k-1}) = \sum_{B \cup S^1 \cup ... \cup S^{n_{k-1}} = X_k} p_{B_k}(B) \prod_{i=1}^{n_k - 1} \pi_{S_{k|k-1}}(S^i | \{x_{k-1}\}) \tag{6}$$

$$\pi_k(S | x_{k-1}^i) = \begin{cases} \mathrm{P}_S(\{x_{k-1}^i\}) \pi_k(\{s\} | x_{k-1}^i), \text{ if } S = s \\ 1 - P_S(\{x_{k-1}^i\}), \text{ if } S = \emptyset \end{cases} \tag{7}$$

D. observation model

The observation model is given by the intersection of clutter measurements and valid measurements :

$$Z_k = C_k \cup \Theta_k \tag{8}$$

A valid association is defined with an association function as follow :

$$\theta_k^i = \begin{cases} j, & \text{if object } i \text{ is associated to measurement } j \\ 0, & \text{if object } i \text{ is undetected} \end{cases} \tag{9}$$

The PDF associated to observation is then :

$$\pi(Z_k|X_k) = \sum_{\theta_k \in \Theta_k} \left(exp(\overline{\lambda_c}) \prod_{j=1}^{m_k} \lambda_c(\{z_k^j\}) \prod_{i:\theta_k^i=0} (1 - P_D(\{x_k^i\}) \prod_{i:\theta_k^i>0} \frac{P_D(\{x_k^i\} g_k(\{z_k^{\theta_k^i}\}|\{x_k^i\})}{\lambda_c(z_k^j)} \right) \tag{10}$$

With $\lambda_c(.)$ the distribution over the clutters generated by the sensors, and $\overline{\lambda_c}$ its expectation. P_D is the probability of an object to be detected ($<= 1$).

The first term of the above equation is linked to clutters, the second to miss detection and the last one to associated objects.

E. Probability Hypothesis Density function

The PHD function, denoted $D_X(x)$, of a RFS X is defined as :

$$D_X(x) = \int p_X(X') \sum_{x' \in X'} \delta(x - x')dX' \tag{11}$$

This function can be seen as the first order moment of a RFS. And The integral of this function over a subspace A of the state space is the expected number of elements in this subspace.

F. The Probability Hypothesis Density Filter

Due to the presence of unknown distributions and of some integrals in the equations above, these equations are not directly computable. Hence, the last step the development of a probability Hypothesis Density (PHD) filter in order to make it computable.

1) The poisson random finite set: Poisson random finite set or poisson point process (PPP) is defined by $\overline{\lambda}$ the number of occurrences in a given interval and $p(.)$, the distribution on the random variables of the random set.

The PDF of a poisson RFS is :

$$D_X(x) = e^{-\overline{\lambda}} \overline{\lambda}^{|X|} p(x_1)...p(x_2) \tag{12}$$

And its PHD is :

$$D_X(x) = \overline{\lambda} p(x) \tag{13}$$

It has been shown that the intensity function of a PPP is :

$$\lambda(x) = \overline{\lambda} p(x) \tag{14}$$

And so the PHD of PPP is its intensity function. For PPP, there is no need no do approximation in order to apply PHD. So most of the time, the challenge will be to approximate the problem with a PPP.

2) PHD filter's equations: The prediction step is :

$$D_{k|k-1}(x|Z^{k-1}) = s_{k|k-1}(x|Z^{k-1}) + b_k(x) \tag{15}$$

With $s_{k|k-1}(x|Z^{k-1})$ being the PHD of the RFS estimating the surviving objects and $b_k(x)$ the PHD of the RFS estimating the newborn objects.

Thus, the steps of PHD algorithm for the prediction will be: predict each hypotheses independently from each others, sum their PHDs, add the newborn phd.

And for update step, the equation is:

$$D_{k|k}(x|Z^k) = D_{k|k-1}(x|Z^{k-1}) \left(1 - P_D(x) + \sum_{z \in Z_k} \frac{P_D(x) g_k(z_k|x)}{\lambda_c(z) + \int P_D(\epsilon) g_k(z_k|\epsilon) D_{k|k-1}(x|Z^{k-1}) d\epsilon} \right) \tag{16}$$

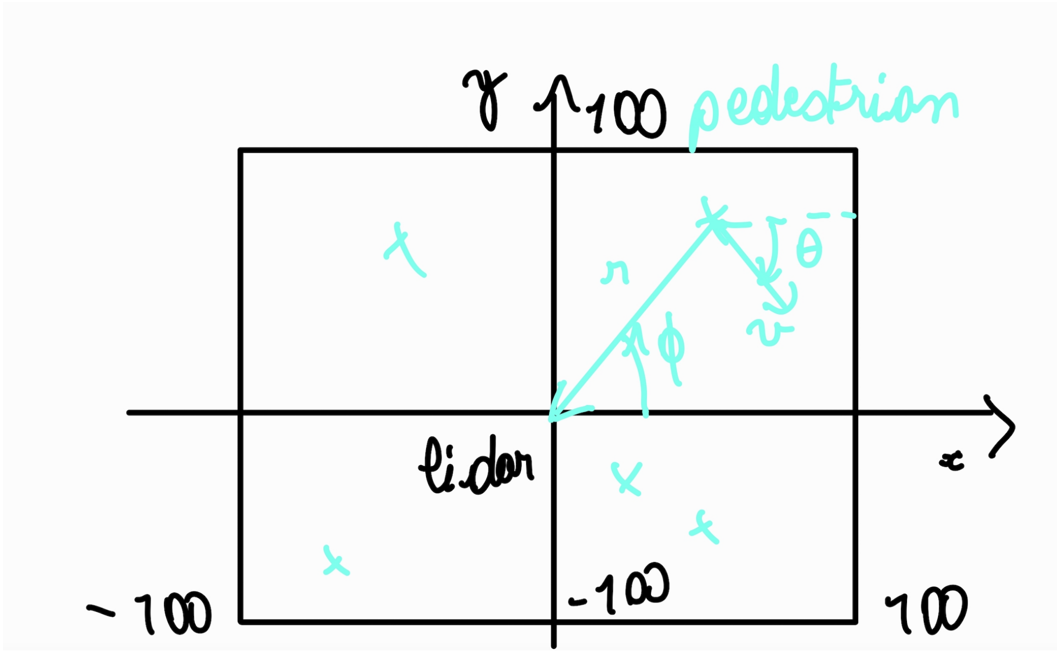

Fig. 1. pedestrian modelisation

II. Application to multi pedestrian tracking

The modelisation of the pedestrian can be seen below 1:

A. Evolution Model

The state representation for the pedestrian could be given as:

$$X_k = \begin{bmatrix} x_k \\ y_k \\ \theta_k \\ v_k \end{bmatrix} \quad X_{k+1} = \begin{bmatrix} x_k + dx \\ y_k + dy \\ \theta_k + d\theta \\ v_k + dv \end{bmatrix} \tag{17}$$

With x, y the Cartesian coordinates, θ the angle formed between the direction of the pedestrian considering the position of the car and the axe x attached to the car. v is the speed of the pedestrian.

The linearized transition model is then given by :

$$X_{k+1} = AX_k \tag{18}$$

With A being a linearization through the Jacobian matrix :

$$A = \begin{bmatrix} \frac{\partial \dot{X}_1}{\partial X_1} & \frac{\partial \dot{X}_1}{\partial X_2} & \frac{\partial \dot{X}_1}{\partial X_3} & \frac{\partial \dot{X}_1}{\partial X_4} \\ \frac{\partial \dot{X}_2}{\partial X_1} & \frac{\partial \dot{X}_2}{\partial X_2} & \frac{\partial \dot{X}_2}{\partial X_3} & \frac{\partial \dot{X}_2}{\partial X_4} \\ \frac{\partial \dot{X}_3}{\partial X_1} & \frac{\partial \dot{X}_3}{\partial X_2} & \frac{\partial \dot{X}_3}{\partial X_3} & \frac{\partial \dot{X}_3}{\partial X_4} \\ \frac{\partial \dot{X}_4}{\partial X_1} & \frac{\partial \dot{X}_4}{\partial X_2} & \frac{\partial \dot{X}_4}{\partial X_3} & \frac{\partial \dot{X}_4}{\partial X_4} \end{bmatrix} \tag{19}$$

Considering the problem, we have the following equations :

$$\begin{cases} x_{k+1} = x_k + vcos(\theta)dt \\ y_{k+1} = y_k + vsin(\theta)dt \end{cases} \tag{20}$$

As we will see later on, there are some assumptions on the speed of the pedestrian and on its direction that imply $v_{k+1} = v_k$ and $\theta_{k+1} = \theta_k$

Given 19, 20 and the above hypothesis, it is possible to write A as :

$$A = \begin{bmatrix} 1 & 0 & -v_k sin(\theta_k)dt & cos(\theta_k)dt \\ 0 & 1 & v_k cos(\theta_k)dt & sin(\theta_k)dt \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{21}$$

B. Observation Model

$$Z_k = \begin{bmatrix} r_k \\ \phi_k \end{bmatrix} \quad = \begin{bmatrix} \sqrt{x_k^2 + y_k^2} \\ arctan(\frac{y_k}{x_k}) \end{bmatrix} \tag{22}$$

The linearized transition model is then given by :

$$Z_k = C X_k \tag{23}$$

With C - considering the equation 22 - equal to :

$$C = \begin{bmatrix} \frac{x_k}{\sqrt{x_k^2+y_k^2}} & \frac{y_k}{\sqrt{x_k^2+y_k^2}} & 0 & 0 \\ \frac{-y_k}{x_k^2+y_k^2} & \frac{x_k}{x_k^2+y_k^2} & 0 & 0 \end{bmatrix} \tag{24}$$

Note that with that observation model, it is impossible to access θ and v.

C. Birth Model

As, θ and v are not directly accessible through the observation variables, we have to directly initialize them. A rough approximation would be $1.31 m/s$ for v which is the average speed of a pedestrian and $\theta = \pi - \phi$ which means that the pedestrian is oriented toward the lidar. This approximation is the only one that can be done because that the movement of the pedestrian is totally unkown. The only assumption that can be done is the fact that as the pedestrian has been detected, he entered the detection area of the lidar (in terms of distance) and so a rough assumption that can be done is the fact that the pedestrian is oriented toward the lidar. θ being the angle $\left(\vec{px}, \vec{pO}\right)$, with p the point representing the pedestrian.

$$x^2 \tan^2(\phi) = y^2 = r^2 - x^2 \tag{25}$$

$$\Rightarrow x^2 \left(1 + tan^2(\phi)\right) = r^2 \tag{26}$$

$$\Rightarrow x = \pm \sqrt{\frac{r^2}{1 + tan^2(\phi)}} \tag{27}$$

$$\tag{28}$$

$$\begin{cases} x = \sqrt{\dfrac{r^2}{1 + tan^2(\phi)}} \quad , \quad if \quad \cos\phi > 0 \\ x = -\sqrt{\dfrac{r^2}{1 + tan^2(\phi)}} \quad , \quad if \quad \cos\phi < 0 \end{cases} \tag{29}$$

As these hypothesis are very inaccurate, we have to put high uncertainties on our birth model. As it is a GM-PHD filter, the birth model is a Gaussian.

D. Log Domain

1) Prediction: ”To avoid overflow and numerical instability due to computation with very small numbers, it is often a good idea to manipulate probabilities in the log domain.”

The equation for the prediction of the weights is given by :

$$w^h_{k|k-1} = P_S w^h_{k-1|k-1} \tag{30}$$

As a remember, the weights intervene inside the compute of the Gaussian mixture PHD (GM-PHD) 31. This model is an approximation of the PHD filter because 15 and 16 are not yet calculable in practice.

$$\sum_{h=1}^{\mathcal{H}_{k|k}} w_{k|k}^{h} \mathcal{N}(x, \mu_{k|k}^{h}, P_{k|k}^{h}) \tag{31}$$

Each Gaussian representing the hypothesis about the presence of a pedestrian or not at a given place. A weight represents the expected number targets under the Gaussian (the confidence about this Gaussian representing an object).

Considering 30, the equation in the log domain will be:

$$l_{k|k-1}^{h} = log(P_S) + l_{k-1|k-1}^{h} \tag{32}$$

2) Update: In the update step, the un-normalized weight $\tilde{w}_{k|k}^{i\mathcal{H}_{k|k-1}+h}$ is given by :

$$\tilde{w}_{k|k}^{i\mathcal{H}_{k|k-1}+h} = P_D w_{k|k-1}^{h} \mathcal{N}(z_k^i | \hat{z}_{k|k-1}^{h}, S_k^h) \tag{33}$$

where $\hat{z}_{k|k-1}^{h}$ is the expected measurement for the hypothesis h and S_k^h the innovation co-variance for this hypothesis. P_D is the probability of an object to be detected

Considering 33, the equation in the log domain will be:

$$\tilde{l}_{k|k}^{i\mathcal{H}_{k|k-1}+h} = log(P_D) + l_{k|k-1}^{h} + log(\mathcal{N}(z_k^i | \hat{z}_{k|k-1}^{h}, S_k^h)) \tag{34}$$

3) Normalization: When all the hypotheses are updated with a given measurement, the corresponding weights are normalized following the next equation:

$$w_{k|k}^{i\mathcal{H}_{k|k-1}+h} = \frac{\tilde{w}_{k|k}^{i\mathcal{H}_{k|k-1}+h}}{\lambda_c(z_k^i) + \sum_{h'=1}^{\mathcal{H}^{k|k-1}} \tilde{w}_{k|k}^{i\mathcal{H}_{k|k-1}+h'}} \tag{35}$$

Considering 35, the equation in the log domain will be:

$$l_{k|k}^{i\mathcal{H}_{k|k-1}+h} = \tilde{l}_{k|k}^{i\mathcal{H}_{k|k-1}+h} - log\left(e^{log(\lambda_c(z_k^i)} + \sum_{h'=1}^{\mathcal{H}^{k|k-1}} e^{\tilde{l}_{k|k}^{i\mathcal{H}_{k|k-1}+h'}}\right) = \tilde{l}_{k|k}^{i\mathcal{H}_{k|k-1}+h} - logsumexp\left(log\left(\lambda_c\left(z_k^i\right)\right), \tilde{l}_{k|k}^{i\mathcal{H}_{k|k-1}+h'}\right) \tag{36}$$