Multi-Target tracking with background discrimination using PHD filters

Roland Jonsson, Johan Degerman, Daniel Svensson, Johannes Wintenby
Electronic Defence Systems,
Saab AB
Göteborg, Sweden
Email: {roland.jonsson,johan.degerman,daniel.x.svensson,johannes.wintenby}@saabgroup.com

***Abstract*—In this paper, we propose a new double PHD filter for simultaneous multi-target tracking and background discrimination for airborne radar applications. Both the foreground and the background processes are modeled as Poisson point processes, which gives a symmetric formulation of the coupled filters. The differences between foreground and background lie in the assumed target dynamics, and in the sensor detection probabilities. Although there are proposals for PHD filter with adaptive background models in the literature, our filter appears to be novel and also the simplest possible. To implement the filter we use a Gaussian mixture approximation of the intensities, which enables simple and effective ways to extract tracks. For the evaluations we use a simulated target tracking scenario with an airborne radar tracking a number of flying targets over a background of road objects. First, the performance of the Gaussian mixture PHD filter with track extraction is illustrated. Second, the superior ability of the foreground-background PHD filter to suppress clutter and disturbing road traffic is illustrated.**

Keywords: Bayesian estimation, target tracking, PHD filter, classification.

I. Introduction

This paper is concerned with multi-target tracking in challenging backgrounds. In particular, we address a scenario where an airborne radar is tracking a number of air targets in a background of a high density of moving ground objects, e.g., road traffic. The classical solution to this problem is to let the radar block detections from a wide radial velocity interval. This is undesirable since it reduces performance and imposes limits on the waveform design for the radar. Another solution would be to track the background objects. However, due to mask effects ground objects are typically seen only for relatively short intervals, and can be very dense in certain areas. Thus, tracking the ground objects individually is unrealistic, and we seek other ways to model the background of ground objects in order to discriminate them from air targets.

In the tracking literature, measurements from non-targets are referred to as clutter, and the concept of a clutter density becomes natural. The most common model the clutter density is by a (non-homogeneous) Poisson point process. Such a clutter model is used in algorithms such as the Probabilistic Data Association (PDA) and the Joint PDA (JPDA) filters [1]. Since the intensity[1] of the Poisson point process is not usually known a priori, various methods to estimate it have been proposed.

Clutter density estimation usually assumes a locally homogeneous intensity, and may use detections from one or more scans (see [2] and references therein). Since a wide variety of phenomena is subsumed in the term clutter, the clutter density could be either constant or totally independent from scan to scan [2]. We note that the problem of road traffic may be somewhere in between those extremes. This leads to the question of how the clutter density evolves over time. This now, is analogous to a recently popular approach to multi target tracking, called the Probability Hypothesis Density (PHD) filter, which models targets as a (non-homogeneous) Poisson point process with an intensity that evolves over time.

There are a couple of works published that address clutter density estimation in PHD filters. In [3], Lian et al. starts with the ordinary PHD formulation that contains an a priori assumed clutter density, which they parameterize as a finite mixture model. They then propose to find the maximum likelihood estimates of parameters describing the clutter through expectation maximization (EM) or Markov Chain Monte Carlo (MCMC) methods. In [4], Chen et al. proposes two different clutter density estimation methods. The first is, like Lian's method, a maximum likelihood estimator of the clutter density in the ordinary PHD formulation. The second method is similar to the one described in this report, in that it evolves two Poisson point processes. It makes some special assumptions on the background process though, and use a Wishart model for the background intensity mixture components. The problem of an unknown background have also been addressed by Mahler, using more complicated models for the clutter process, and mainly directed towards the Cardinalized PHD (CPHD) filter [5]–[7].

In this paper, we propose a new double PHD filter, which is symmetric in its formulation of a foreground and a background model as Poisson point processes. The difference between the foreground and background filters lie in the parameter settings for the dynamics, and in the detection probabilities. It is assumed that each detection either comes from a point

[1]Here intensity refers to the intensity in point process theory, not the clutter signal strength. In target tracking literature the intensity of the clutter process is often referred to as clutter measurement density.

in the foreground or from the background process, and that those two processes are independent. To implement the filter, we use a Gaussian mixture approximation of the intensities as described by Vo et al. [8]. The double PHD filter is evaluated on a target tracking scenario, with an airborne radar tracking a number of flying targets over a background of ground clutter and detections from dense road traffic. The simulations demonstrate a superior ability of the double PHD filter to suppress the clutter and disturbing road traffic, compared to the ordinary PHD filter with track extraction.

II. Problem Formulation

In the PHD filter framework, the total state of all the targets are seen as a Random Finite Set (RFS). At a time *k* the state is

$$X_k = \left\{x_{k,1}, \dots x_{k,M(k)}\right\} \tag{1}$$

of all the states $x_{k,m}$ of all the $M(k)$ individual target states. Note, that there is no target-ID or ordering present. Likewise, the measurements at time *k* are collected in the set

$$Z_k = \left\{z_{k,1}, \dots z_{k,N(k)}\right\}. \tag{2}$$

The optimal solution to the target tracking problem is then to calculate the RFS density $p(X_k|Z_k)$, and to derive optimal estimates of the individual target states from it. Such an approach is infeasible, and the idea of the PHD filter is to recursively estimate the first moment of the RFS X_k, referred to as the Probability Hypothesis Density (PHD) [9], [10].

The problem studied in this paper is to perform target tracking under disturbing background clutter generated by ground-based traffic, and random ground clutter. The problem is thus to discriminate between the targets of interest—called the foreground targets—and the background objects and clutter.

The dynamics of a general target is modeled by a linear-Gaussian process model

$$f_{k|k-1}(x|\xi) = \mathcal{N}(x|F\xi, Q), \tag{3}$$

where ξ is the state at time $k-1$. The probability that a target still exists at the updated time step, given that it had state x in the previous time step is denoted $p_S(x)$. The appearance of new targets is modeled as another independent Poisson point process with the ("birth") intensity $\gamma(x)$. Note that we could have made $p_S(x)$ and $\gamma(x)$ dependent on the time k, and introduced a spawning model, but we will not use that in this paper.

The measurements are modeled through a detection probability $p_D(x)$ (we ignore a possible time dependence here too), and a likelihood $g_k(z|x)$ of a target with state x being the source of measurement z. The considered measurement model is non-linear:

$$z_k = h(x_k) + \varepsilon_k, \tag{4}$$

where ε_k is zero mean Gaussian with covariance matrix R. Measurements can also come from clutter, which is modeled as another independent Poisson point process (on the measurement space) with intensity $\kappa_k(z)$, and from ground moving objects on roads.

III. The Gaussian-mixture PHD filter

The mathematics of random finite sets or equivalently "simple point processes" is quite advanced. In the original derivation of the PHD filter, finite-set statistics (FISST) was used. We will not use that here, but will instead use a conventional measure formalism of probability (with densities defined with respect to unit rate Poisson point process)[2].

The PHD filter models the multi-target state X_k as a Poisson point process, which is fully described by its intensity $\nu_k(x)$. If $\nu_k(x)$ is integrated over some region of the state space, one gets the expected number of targets in that region. Moreover, the number of targets in that region is Poisson distributed with that mean. With this notation, and the notation in Section II, the PHD filter has the form

$$\nu_{k|k-1}(x) = \int p_S(\xi) f_{k|k-1}(x|\xi)\nu_{k-1}(x)d\xi + \gamma(x) \tag{5}$$

$$\begin{aligned}\nu_k(x) = (1 - p_D(x))\, \nu_{k|k-1}(x) + \dots \\ \sum_{z \in Z_k} \frac{p_D(x) g_k(z|x)\nu_{k|k-1}(x)}{\kappa_k(z) + \int p_D(x) g_k(z|x) \nu_{k|k-1}(x)dx}\end{aligned} \tag{6}$$

The expression above contains an approximation, since the a posteriori point process is no longer Poisson, but is approximated with a Poisson point process with the same intensity.

Practical implementations of the above PHD expressions require approximations. Two such approximations are given by particle-filter implementations [11], which require large number of particles and some kind of clustering technique to extract tracks, and Gaussian-sum approximations of the intensities, as described by Vo et al. [8]. In this paper, we use Gaussian-mixture implementations. We thus represent the intensity by:

$$\nu_k(x) = \sum_{i=1}^{J_k} w_k^{(i)} \mathcal{N}(x|m_k^{(i)}, P_k^{(i)})$$

where $\mathcal{N}(x|m, P)$ denotes a Gaussian density with mean m and covariance matrix P. We recognize that, if all other quantities depending on x present in the PHD expression also are represented as Gaussian sums, the expression for $\nu_k(x)$ will hold for all k with an exponentially increasing number of components J_k. In practice, this increase in components has to be countered by some pruning and merging technique. The details of how the various parameters are updated can be found in [8] and we will not write them out here.

IV. A PHD filter estimating both foreground and background models

In this section, we will formulate a double PHD filter for estimating both foreground targets and background clutter. It is assumed that each detection either comes from a point in the foreground process, or from the background process, and that those two processes are independent. In Appendix A, it is shown that such a filter has the form:

[2]For the relation to the FISST formalism we refer to [11]

$$\nu^{\mathrm{F}}_{k|k-1}(x) = \int p^{\mathrm{F}}_S f^{\mathrm{F}}_{k|k-1}(x|\xi)\nu^{\mathrm{F}}_{k-1}(x)d\xi + \gamma^{\mathrm{F}}(x)$$

$$\nu^{\mathrm{B}}_{k|k-1}(x) = \int p^{\mathrm{B}}_S f^{\mathrm{B}}_{k|k-1}(x|\xi)\nu^{\mathrm{B}}_{k-1}(x)d\xi + \gamma^{\mathrm{B}}(x)$$

$$\nu^{\mathrm{F}}_k(x) = \left(1 - p^{\mathrm{F}}_D\right)\nu^{\mathrm{F}}_{k|k-1}(x) + \dots$$
$$\sum_{z\in Z_k} \frac{p^{\mathrm{F}}_D g_k(z|x)\nu^{\mathrm{F}}_{k|k-1}(x)}{\int p^{\mathrm{F}}_D g_k(z|x)\nu^{\mathrm{F}}_{k|k-1}(x)dx + \int p^{\mathrm{B}}_D g_k(z|x)\nu^{\mathrm{B}}_{k|k-1}(x)dx}$$

$$\nu^{\mathrm{B}}_k(x) = \left(1 - p^{\mathrm{B}}_D\right)\nu^{\mathrm{B}}_{k|k-1}(x) + \dots$$
$$\sum_{z\in Z_k} \frac{p^{\mathrm{B}}_D g_k(z|x)\nu^{\mathrm{B}}_{k|k-1}(x)}{\int p^{\mathrm{F}}_D g_k(z|x)\nu^{\mathrm{F}}_{k|k-1}(x)dx + \int p^{\mathrm{B}}_D g_k(z|x)\nu^{\mathrm{B}}_{k|k-1}(x)dx}$$

and use a Gaussian sum approximation of the intensities:

$$\nu^{\mathrm{F}}_k(x) = \sum_{i=1}^{J^{\mathrm{F}}_k} F^{\mathrm{F}} w^{\mathrm{F}(i)}_k \mathcal{N}(x|m^{\mathrm{F}(i)}_k, P^{\mathrm{F}(i)}_k)$$

$$\nu^{\mathrm{B}}_k(x) = \sum_{i=1}^{J^{\mathrm{B}}_k} F^{\mathrm{B}} w^{\mathrm{B}(i)}_k \mathcal{N}(x|m^{\mathrm{F}(i)}_k, P^{\mathrm{B}(i)}_k)$$

where the $^{\mathrm{F}}$ and $^{\mathrm{B}}$ indeces denote foreground and background PHD filter respectively. Note that we use the same measurement model for both the foreground and the background, which need to be linearized in EKF fashion as described in the subsequent section. The differentiation between foreground from background is based on choosing: $f^{\mathrm{F}}_{k|k-1}(x|\xi) = \mathcal{N}(x|F^{\mathrm{F}}\xi, Q^{\mathrm{F}})$, $f^{\mathrm{B}}_{k|k-1}(x|\xi) = \mathcal{N}(x|F^{\mathrm{B}}\xi, Q^{\mathrm{B}})$, and in the birth intensities $\gamma^{\mathrm{F}}(x) = w^{\mathrm{F}}_\gamma \mathcal{N}(x|m^{\mathrm{F}}_\gamma, P^{\mathrm{F}}_\gamma)$ and $\gamma^{\mathrm{B}}(x) = w^{\mathrm{B}}_\gamma \mathcal{N}(x|m^{\mathrm{B}}_\gamma, P^{\mathrm{B}}_\gamma)$. The dynamics for the background should be chosen to reflect that the background objects are "stationary", even for moving objects like road traffic, in the sense that they are localized in certain areas. We give an example of this in the next section, where some simulation results are presented.

Another difference we make between the processing of the foreground and the background model, is that another merging algorithm is used for the background model. The algorithm in [8] used for the foreground model is not suitable for the background, since it favors strong components. For the background model we instead use the algorithm by Runnalls [12].

V. Implementation strategies and track extraction

To handle the non-linear measurement model, we use a Extended Kalman Filter (EKF), like the extension to the PHD filter described in [8]. The problem is to choose the point to linearize around. In [8], the updated means m_{k-1} of the components are used as in the EKF. This works well for all old components, but not for the single birth intensity component. For the new components, we instead linearize around a projection $\hat{x}(z)$ of the measurement (i.e. $h(\hat{x}(z)) = z$). Note also that for the new components, it is numerically better to compute $m^{(i)}_k$ and $P^{(i)}_k$ directly, and not to use the Kalman gain and adjustments to m_γ and P_γ.

To reduce the number of components, we remove components with $w^{(i)}_k < T$ where the threshold T typically is some small value like 0.01. Similar components are merged using the algorithm in [8]. Here one successively looks at the component with the largest weight $w^{(i)}_k$, and merge together all components such as:

$$\left(m^{(i)}_k - m^{(j)}_k\right)^T \left(P^{(i)}_k\right)^{-1} \left(m^{(i)}_k - m^{(j)}_k\right) < U.$$

When using a PHD filter for tracking, an additional procedure to extract estimates for individual targets from $\nu_k(x)$, and also to form tracks where the identity of a target is preserved over time. This is a problem where there is no strong theoretical basis. Ad hoc methods appear to perform relatively well though. In practice, we are interested in the performance after some track extraction procedure, so the number of targets estimated from $\nu_k(x)$ at a certain k is not in itself of interest. It was proposed in [8] to use the components (after the pruning and merging) of $\nu_k(x)$ with $w^{(i)}_k$ over some threshold as the target estimates, regardless of the mean number of targets given by $\sum_i w^{(i)}_k$. This was in [13] combined with a track extraction procedure. We use a simplified version of those ideas.

We introduce an ID-tag on each component of the Gaussian sum. This tag, for new components, gets a new value for each detection. For components that are the result of the updating of old components, the old tag is retained. When components are merged, the oldest tag is retained. We also keep track of the age of a component, in number of updates. A new track is started when there is a component with $w^{(i)}_k > w_{track}$ and an age larger than a_{track}. We have typically used $w_{track} = 0.9$ and $a_{track} = 3$. When a track is started, it is marked by the ID-tag of the component that started it, and all components with that tag will thereafter be considered a part of the track. When a component is part of a track, we lower the pruning threshold T to $T/10$ in order not to lose tracks too often.

VI. Simulation results

In this section we will present some results from a target tracking simulation. Our sensor is an airborne radar, tracking a number of flying targets, and being disturbed by traffic on roads. The trajectories of the targets are presented in Figure 1. We will start by examining the performance of the Gaussian sum PHD filter with track extraction on random non-correlated uniformly distributed clutter detections. After that, the foreground-background filter is introduced and its ability to suppress uniform clutter and disturbing road traffic is illustrated.

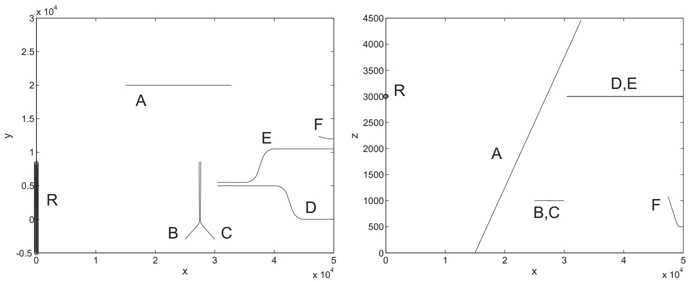

Figure 1. Target scenario. The radar (R) moves at an altitude of 3000 m along x=0, from y=-5 km to y=8 km. One target (A) moves along y=20 km while climbing from 0 to 4500 m. Two targets (B and C) at an altitude of 1000 m moves towards each other, crosses each others path, and then moves with close separation in the y direction. Two other targets (D and E), at an altitude of 3000 m, moves in the -x direction from x=50 km, manoeuvre and end close together. Finally, a slowly moving target (F) starts at an altitude of 500 m from x=50 km and starts climbing.

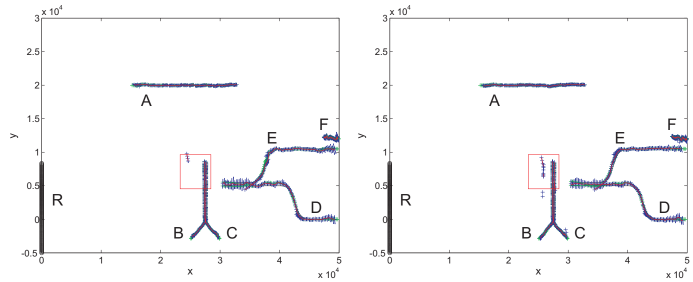

Figure 2. Conventional PHD. Clutter from a limited ground area. The clutter density is 0.1 per km^2 in the left panel and 0.2 per km^2 in the right panel. $\kappa = 10^{-9}$.

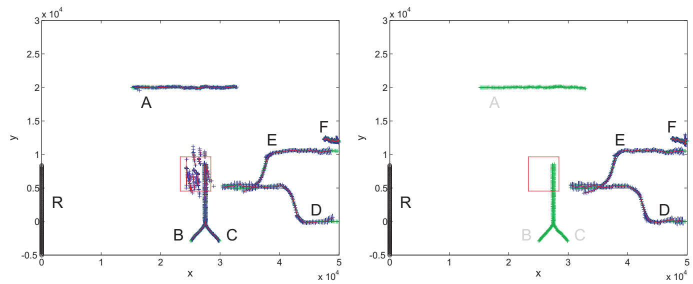

Figure 3. Conventional PHD. Clutter from a limited ground area. The clutter density is 0.5 per km^2. $\kappa = 10^{-9}$ in the left panel and $\kappa = 10^{-8}$ in the right panel.

We are using a state with positions and velocities, $\mathbf{x}_k = [x_k, y_k, z_k, \dot{x_k}, \dot{y_k}, \dot{z_k}]$, and a motion model

$$\mathbf{x}_k = \mathbf{F}\mathbf{x}_{k-1} + \mathbf{v}$$

where $\mathbf{v}$ have covariance matrix $\mathbf{Q}^{\mathrm{F}}$, and

$$\mathbf{F}^{\mathrm{F}} = \begin{bmatrix} \mathbf{I}_3 & T_0\mathbf{I}_3 \\ \mathbf{0} & \mathbf{I}_3 \end{bmatrix} \qquad \mathbf{Q}^{\mathrm{F}} = q^{\mathrm{F}} \begin{bmatrix} \frac{T_0^3}{3}\mathbf{I}_3 & \frac{T_0^2}{2}\mathbf{I}_3 \\ \frac{T_0^2}{2}\mathbf{I}_3 & T_0\mathbf{I}_3 \end{bmatrix}$$

We have the time between measurements $T_0 = 1$ and use $q^{\mathrm{F}} = 50$ for the system noise. Measurements are in polar coordinates and Doppler velocity: $\mathbf{z}_k = [r_k, \varphi_k, \theta_k, \dot{r_k}]$. The measurement model thus becomes:

$$\mathbf{z}_k = h(\mathbf{x}_k - \mathbf{x}_0) + \mathbf{w} \quad h(\mathbf{x}) = \begin{bmatrix} \sqrt{x^2+y^2+z^2} \\ \tan^{-1}(y/x) \\ \tan^{-1}\left(z/\sqrt{x^2+y^2}\right) \\ \frac{x\dot{x}+y\dot{y}+z\dot{z}}{\sqrt{x^2+y^2+z^2}} \end{bmatrix}$$

where $\mathbf{w}$ have covariance matrix $\mathbf{R} = diag(\left[\sigma_r^2, \sigma_\varphi^2, \sigma_\theta^2, \sigma_{\dot{r}}^2\right])$ and $\mathbf{x}_0$ is the position of the radar. We used $\sigma_r = 10$, $\sigma_\varphi = \sigma_\theta = 0.005$ and $\sigma_{\dot{r}} = 2$. Measurements were simulated with errors according to $\mathbf{R}$ and with a 0.9 probability of detection.

The PHD filter used: $p_D^{\mathrm{F}} = 0.85$, $p_S^{\mathrm{F}} = 0$ and $\kappa^{\mathrm{F}} = 0$. The birth intensity was set to $w_\gamma^{\mathrm{F}} = 0.1$, and with m_γ^{F} at the center of the scene, at the same altitude as the radar and at zero velocity. The matrix P_γ^{F} was set to $P_\gamma = diag(\left[10^{10}, 10^{10}, 25\cdot 10^6, 25\cdot 10^4, 25\cdot 10^4, 25\cdot 10^4\right]$.

We will now introduce clutter from the ground, by introducing random detections in a 4 x 5 km rectangle at x=24 km and y=5km, near target B and C. The density of clutter is varied as 0.1, 0.2 and 0.5 detections per km^2. The clutter intensity parameter in PHD filter was set to $\kappa^{\mathrm{B}} = 10^{-9}$. The results are shown in figure 2 and 3. We note that the number of false tracks increase from a single one at 0.1 detections per km^2 to about 25 at 0.5 detections per km^2. The sensitivity to the clutter intensity parameter is also quite high. When increased to $\kappa^{\mathrm{B}} = 10^{-8}$, the tracks of target A, B, and C disappeared along with all the clutter. This illustrates the difficulty in handling clutter using the traditional PHD filter, and we will now switch to our dual PHD filter, with both foreground and background models.

In using the PHD filter with both foreground and background models as described in section IV we use as foreground exactly the same parameter settings as above. For the background we use

$$\mathbf{F}^{\mathrm{B}} = \begin{bmatrix} \mathbf{I}_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \qquad \mathbf{Q}^{\mathrm{B}} = Diag\left([500, 500, 0, 10, 10, 0]\right)$$

We thus have clutter points that can "move" like a random walk over the ground, but their velocity distribution is fixed around zero and there is no movement in the stationary z direction. The filter parameters used: $P_D^{\mathrm{B}} = 0.3$, $p_S^{\mathrm{B}} = 0$, $w_\gamma^{\mathrm{B}} = 3$. We also used m_γ^{B} at the center of the scene as before, but at zero altitude. The matrix P_γ^{B} was set to $P_\gamma^{\mathrm{B}} = diag(\left[10^{10}, 10^{10}, 100, 100, 100, 0.01\right]$. The threshold for the background merging algorithm was set to 20.

When applying this foreground-background PHD filter on same clutter case as the last one above, we get the results shown in figure 4. In this figure we have included the intensity of the background model at the last time step, plotted in magenta. One notes that a number of Gaussian mixture components cover the area of the clutter, and that no false tracks were generated. Increasing the clutter density to 5 per km^{2}also worked well and the result is shown in figure 5. Moving the clutter area out to the low flying slow target F at x=50 km will result in losing that track at the high clutter density of 5 per km^2, but at the more moderate 0.5 per km^2 the target F is tracked as shown in figure 5.

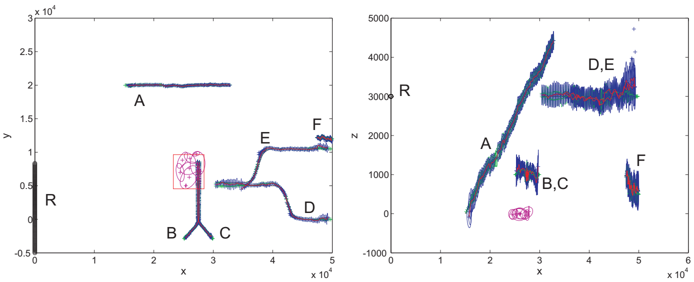

Figure 4. Foreground-background PHD. Clutter from a limited ground area. The clutter density is 0.5 per km^2.

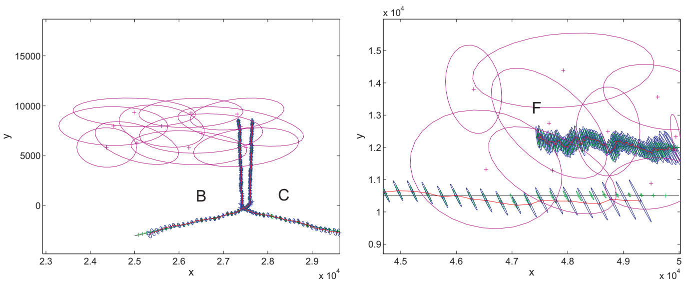

Figure 5. Foreground-background PHD. Clutter from a limited ground area. In the left panel the clutter density is 5 per km^2, and in the right panel the clutter density is 0.5 per km^2, located near low-flying target F.

We are now turning our attention to the more challenging problem of interfering road traffic. We will use the same target scenario as above, and replace clutter with simulated road traffic. To generate the road traffic detections we used a tool that, using a geographic database, can compute the visibility of targets moving on roads. We picked out a few roads, and simulated new vehicles starting with 0.02 probability each second. Those vehicles are traveling along the roads with a speed that is picked randomly in an $\pm 30\%$ interval around the speed limit for the road. Detections where generated with 0.3 probability if the Doppler speed exceeded 10 m/s. An example of detections at a certain time step superimposed on the target trajectories are given in figure 6.

Running the conventional PHD filter on the data with road traffic results in about 500 short tracks, and is illustrated in figure 7. Using the foreground-background PHD filter on the other hand, results in only a single short track from a road, as can be seen in figure 8. In order to assess the importance of altitude separation between the targets and the road traffic we let the birth intensity parameters for the background be the same as for the foreground, i.e. $m_\gamma^{\mathrm{B}} = m_\gamma^{\mathrm{F}}$ and $P_\gamma^{\mathrm{B}} = P_\gamma^{\mathrm{F}}$. In that case the number of vehicle tracks increased to 9, and the results are illustrated in figure 9. The most problematic case, in this example, is to discriminate the low flying target at x=50 km from vehicle tracks. This is illustrated in figure 10. The experiment indicate that, while the difference in dynamics between the foreground and the background is important for the suppression of road traffic, the initiation of new tracks is sensitive and separation of birth intensities $(w_\gamma^{\mathrm{B}} \gg w_\gamma^{\mathrm{F}})$ highly beneficial.

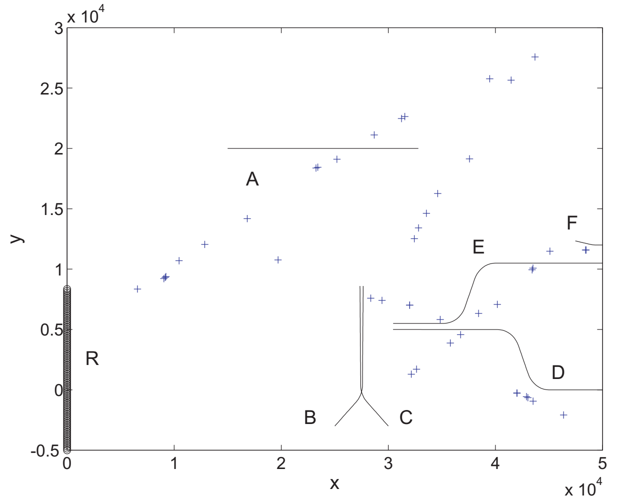

Figure 6. Example of detections of road traffic at a certain time step.

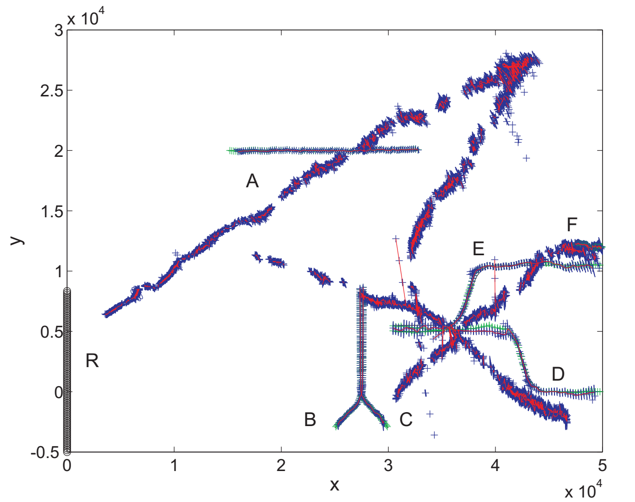

Figure 7. Conventional PHD filter with road traffic.

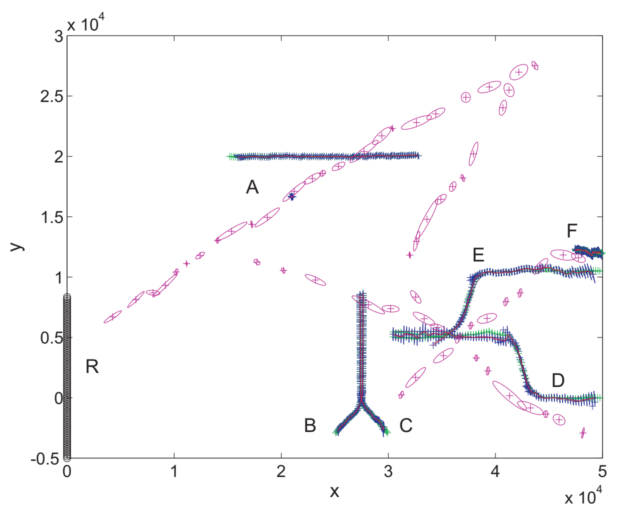

Figure 8. Foreground-background PHD filter with road traffic.

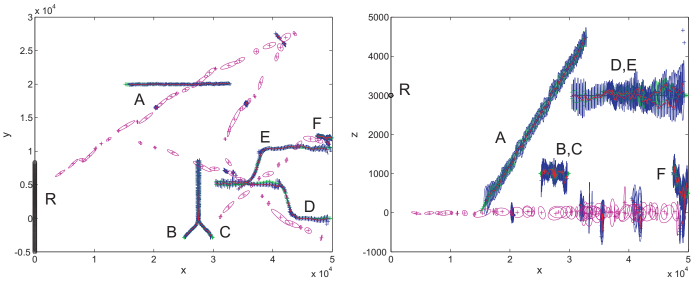

Figure 9. Foreground-background PHD filter with road traffic. Birth intensity parameters for the background set to the same as for the foreground.

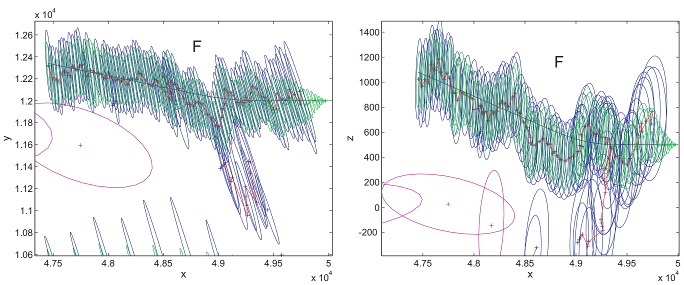

Figure 10. Foreground-background PHD filter with road traffic. Birth intensity parameters for the background set to the same as for the foreground. Detail of low flying target (F).

VII. Conclusions

We have implemented a Gaussian mixture PHD filter together with a simple track extraction procedure. Using this filter in a simulation of an airborne radar, tracking a number of flying targets, showed good results.

Adding simulated clutter and road traffic overwhelmed the PHD filter. Although, there is a clutter intensity parameter in the PHD filter, it is sensitive and hard to tune. In the case of road traffic, hundreds of short tracks where started.

To handle challenging backgrounds, such as road traffic, we propose a new double PHD filter that is symmetric in its formulation of a foreground and a background model as Poisson point processes. The foreground is differentiated from the background using different parameter settings for the dynamics and detection probabilities. It is assumed that each detection either comes from a point in the foreground or the background process, and that those two processes are independent. Although there is proposals for PHD filter with adaptive background models in the literature, our filter appears to be novel and also the simplest possible. Testing our new filter on the same airborne radar scenario, with simulated clutter and road traffic, showed excellent performance. Very high clutter intensities were handled perfectly, and road traffic was also mitigated.

Appendix A
Derivation of the Double PHD filter

The idea of the Double PHD filter is to propagate two different, and *independent*, Poisson PP:s $\mathbf{X}$ and $\mathbf{Y}$ represented by their intensities $\nu_{\mathbf{X}}(x)$ and $\nu_{\mathbf{Y}}(y)$. The prediction part of the filter is separate for $\mathbf{X}$ and $\mathbf{Y}$ and is equivalent to the ordinary PHD filter. The different dynamics for $\mathbf{X}$ and $\mathbf{Y}$, e.g. for targets and some background objects, are thus reflected in different parameters in the prediction step.

In the measurement updating part of the filter, the process $\mathbf{Y}$ takes the place of clutter process $\mathbf{K}$ in he ordinary PHD filter. This results in the updating rule:

$$\nu_{\mathbf{X}}(x) \quad \leftarrow \quad (1 - p_{D\mathbf{X}}(x))\,\nu_{\mathbf{X}}(x) + \ldots$$
$$\sum_{z \in Z} \frac{p_{D\mathbf{X}}(x) l_{\mathbf{X}}(z|x) \nu_{\mathbf{X}}(x)}{\int p_{D\mathbf{Y}}(y) l_{\mathbf{Y}}(z|y) \nu_{\mathbf{Y}}(y) dy + \int p_{D\mathbf{X}}(x) l_{\mathbf{X}}(z|x) \nu_{\mathbf{X}}(x) dx}$$

$$\nu_{\mathbf{Y}}(y) \quad \leftarrow \quad (1 - p_{D\mathbf{Y}}(y))\,\nu_{\mathbf{Y}}(y) + \ldots$$
$$\sum_{z \in Z} \frac{p_{D\mathbf{Y}}(y) l_{\mathbf{Y}}(z|y) \nu_{\mathbf{Y}}(y)}{\int p_{D\mathbf{Y}}(y) l_{\mathbf{Y}}(z|y) \nu_{\mathbf{Y}}(y) dy + \int p_{D\mathbf{X}}(x) l_{\mathbf{X}}(z|x) \nu_{\mathbf{X}}(x) dx}$$

which we will prove below.

The detection procedure splits $\mathbf{X}$ in two independent PP:s $\mathbf{X} = \check{\mathbf{X}} \cup \hat{\mathbf{X}}$. The undetected part $\check{\mathbf{X}}$ is Poisson PP with intensity $(1 - p_{D\mathbf{X}}(x))\,\nu_{\mathbf{X}}(x)$ and the detected part $\hat{\mathbf{X}}$ is Poisson PP with intensity $p_{D\mathbf{X}}(x)\nu_{\mathbf{X}}(x)$. Equivalently for $\mathbf{Y} = \check{\mathbf{Y}} \cup \hat{\mathbf{Y}}$ with intensities $(1 - p_{D\mathbf{Y}}(y))\,\nu_{\mathbf{Y}}(y)$ and $p_{D\mathbf{Y}}(y)\nu_{\mathbf{Y}}(y)$. The measurement process $\mathbf{Z}$ is a superposition of $\mathbf{Z}|\mathbf{X}$ and $\mathbf{Z}|\mathbf{Y}$, which are two independent Poisson PP:s. Using probability generating functionals we thus have

$$G_{\mathbf{Z}|\mathbf{X},\mathbf{K}}\left[h|z_1,\ldots z_n\right] = G_{\mathbf{Z}|\hat{\mathbf{X}}}\left[h|z_1,\ldots z_n\right] G_{\mathbf{Z}|\hat{\mathbf{Y}}}\left[h|z_1,\ldots z_n\right]$$

The joint probability generating functional, due to the independence of $\mathbf{X}$ and $\mathbf{Y}$, now becomes:

$$G_{\hat{\mathbf{X}},\hat{\mathbf{Y}},\mathbf{Z}}\left[f,g,h\right] = G_{\hat{\mathbf{X}}}\left[f H_{\mathbf{X}}\left[h|.\right]\right] G_{\hat{\mathbf{Y}}}\left[g H_{\mathbf{Y}}\left[h|.\right]\right]$$

$$H_{\mathbf{X}}\left[h|x\right] = \int h(z) l_{\mathbf{X}}(z|x) dz \qquad H_{\mathbf{Y}}\left[h|y\right] = \int h(z) l_{\mathbf{Y}}(z|y) dz$$

and

$$G_{\hat{\mathbf{X}}}\left[g\right] = \exp\left(\int (g(x) - 1)\, p_{D\mathbf{X}}(x) \nu_{\mathbf{X}}(x) dx\right)$$

$$G_{\hat{\mathbf{Y}}}\left[g\right] = \exp\left(\int (g(y) - 1)\, p_{D\mathbf{Y}}(y) \nu_{\mathbf{Y}}(y) dy\right)$$

The functional derivatives of $G_{\hat{\mathbf{X}},\hat{\mathbf{Y}},\mathbf{Z}}\left[f,g,h\right]$ with respect to h becomes

$$\frac{\delta^n}{\delta h(z_1)\ldots\delta h(z_n)} G_{\hat{\mathbf{X}},\hat{\mathbf{Y}},\mathbf{Z}}\left[f,g,h\right] =$$
$$G_{\hat{\mathbf{X}},\hat{\mathbf{Y}},\mathbf{Z}}\left[f,g,h\right] \prod_{i=1}^{n} \Big(\int f(x) l_{\mathbf{X}}(z_i|x) p_{D\mathbf{X}}(x) \nu_{\mathbf{X}}(x) dx +$$
$$\int g(y) l_{\mathbf{Y}}(z_i|y) p_{D\mathbf{Y}}(y) \nu_{\mathbf{Y}}(y) dy \Big)$$

At this stage, we need Bayes formula for probability generating functionals, which is written as

$$G_{\mathbf{X},\mathbf{Y}|\mathbf{Z}}[f,g|y_1,\dots y_n] = \frac{\frac{\delta^n}{\delta h(y_1)\dots\delta h(y_n)} G_{\mathbf{X},\mathbf{Y},\mathbf{Z}}[f,g,h]\Big|_{h=0}}{\frac{\delta^n}{\delta h(y_1)\dots\delta h(y_n)} G_{\mathbf{X},\mathbf{Y},\mathbf{Z}}[1,1,h]\Big|_{h=0}}$$

Its derivation is lengthy and requires a background in point process theory which cannot be given here. Using this equation, and $G_{\hat{\mathbf{X}},\hat{\mathbf{Y}},\mathbf{Z}}[f,g,0] = G_{\hat{\mathbf{X}}}[0]\, G_{\hat{\mathbf{Y}}}[0]$, we have

$$G_{\hat{\mathbf{X}},\hat{\mathbf{Y}}|\mathbf{Z}}[f,g|z_1\dots z_n] = \prod_{i=1}^{n} \frac{\int f(x) l_{\mathbf{X}}(z_i|x) p_{D\mathbf{X}}(x)\nu_{\mathbf{X}}(x)dx + \dots}{\int l(z_i|x) p_{D\mathbf{X}}(x)\nu_{\mathbf{X}}(x)dx + \dots} \frac{\dots \int g(y) l_{\mathbf{Y}}(z_i|y) p_{D\mathbf{Y}}(y)\nu_{\mathbf{Y}}(y)dy}{\dots \int l_{\mathbf{Y}}(z_i|y) p_{D\mathbf{Y}}(y)\nu_{\mathbf{Y}}(y)dy}$$

In the ordinary PHD filter the approximation of the (a posteriori) point process with a Poisson process with same intensity can bee seen as choosing the Poisson PP $\tilde{\mathbf{X}}$ which minimizes the Kullback Leibler divergence:

$$\int \log \frac{dP_{\mathbf{X}}}{dP_{\tilde{\mathbf{X}}}} dP_{\mathbf{X}}$$

or equivalently maximizes:

$$\int \log\left(\pi_{\tilde{\mathbf{X}}}(\mathbf{x})\right) dP_{\mathbf{X}}$$

This is achieved by choosing $\nu_{\tilde{\mathbf{X}}}(x) = \nu_{\mathbf{X}}(x)$, which is proved in lemma A.3 in [14]. Now if we want to approximate the bivariate process $\mathbf{X}$,$\mathbf{Y}$ with two independent Poisson PP $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{Y}}$, we must maximize

$$\int \log\left(\pi_{\tilde{\mathbf{X}}}(\mathbf{x})\right) dP_{\mathbf{X},\mathbf{Y}} + \int \log\left(\pi_{\tilde{\mathbf{Y}}}(\mathbf{x})\right) dP_{\mathbf{X},\mathbf{Y}}$$

which is achieved by choosing the intensities of $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{Y}}$ as respective marginal intensity of the $\mathbf{X}$,$\mathbf{Y}$ process. It is quite easy to show, from the definition of the joint probability generating functional, that

$$\nu_{\mathbf{X}}(x) = \frac{\delta G_{\mathbf{X},\mathbf{Y}}[f,g]}{\delta f(x)}\bigg|_{f=1,g=1}$$

$$\nu_{\mathbf{Y}}(y) = \frac{\delta G_{\mathbf{X},\mathbf{Y}}[f,g]}{\delta g(y)}\bigg|_{f=1,g=1}$$

Now we can use the expression for $G_{\hat{\mathbf{X}},\hat{\mathbf{Y}}|\mathbf{Z}}[f,g|z_1\dots z_n]$ to find that:

$$\nu_{\hat{\mathbf{X}}|\mathbf{Z}}(x) = \sum_{i=1}^{n} \frac{l_{\mathbf{X}}(z_i|x) p_{D\mathbf{X}}(x)\nu_{\mathbf{X}}(x)}{\int l(z_i|x) p_{D\mathbf{X}}(x)\nu_{\mathbf{X}}(x)dx + \int l_{\mathbf{Y}}(z_i|y) p_{D\mathbf{Y}}(y)\nu_{\mathbf{Y}}(y)dy}$$

$$\nu_{\hat{\mathbf{Y}}|\mathbf{Z}}(y) = \sum_{i=1}^{n} \frac{l_{\mathbf{Y}}(z_i|y) p_{D\mathbf{Y}}(y)\nu_{\mathbf{Y}}(y)}{\int l(z_i|x) p_{D\mathbf{X}}(x)\nu_{\mathbf{X}}(x)dx + \int l_{\mathbf{Y}}(z_i|y) p_{D\mathbf{Y}}(y)\nu_{\mathbf{Y}}(y)dy}$$

which combined with the intensities for the undetected points $\check{\mathbf{X}}$ and $\check{\mathbf{Y}}$ results in the expressions in the updating rule.

References

[1] S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*. Artech House, 1999.

[2] T. L. Song and D. Musicki, "Adaptive clutter measurement density estimation for improved target tracking," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 2, pp. 1457–1466, apr 2011.

[3] F. Lian, C. Han, and W. Liu, "Estimating unknown clutter intensity for PHD filter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 46, no. 4, pp. 2066–2078, oct 2010.

[4] X. Chen, R. Tharmarasa, T. Kirubarajan, and M. Pelletier, "Integrated clutter estimation and target tracking using poisson point process," in *Proceedings of SPIE Volume: 7445*, 2009.

[5] R. Mahler, "CPHD and PHD filters for unknown backgrounds II: multitarget filtering in dynamic clutter," in *Proceedings of SPIE Volume: 7330*, 2009.

[6] R. Mahler and A. El-Fallah, "CPHD and PHD filters for unknown backgrounds III: Tractable multitarget filtering in dynamic clutter," in *Proceedings of SPIE Volume: 7698*, 2010.

[7] R. P. S. Mahler, B. T. Vo, and B. N. Vo, "CPHD filtering with unknown clutter rate and detection profile," *Accepterd for publication in IEEE Transactions on Signal Processing*, 2011.

[8] B.-N. Vo and W.-K. Ma, "The gaussian mixture probability hypothesis density filter," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4091–4104, nov 2006.

[9] R. P. S. Mahler, "Multitarget bayes filtering via first-order multitarget moments," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 39, no. 4, pp. 1152–1178, oct 2003.

[10] ——, *Statistical multisource-multitarget information fusion*. Artech House, 2007.

[11] B.-N. Vo, S. Singh, and A. Doucet, "Sequential Monte Carlo methods for multi-target filtering with random finite sets," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 41, no. 4, pp. 1224–1245, oct 2005.

[12] A. R. Runnalls, "Kullback-Leibler approach to Gaussian mixture reduction," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 3, pp. 989–999, jul 2007.

[13] K. Panta, D. E. Clark, and B.-N. Vo, "Data association and track management for the gaussian mixture probability hypothesis density filter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 45, no. 3, pp. 1003–1016, jul 2009.

[14] S. S. Singh, B.-N. Vo, A. Baddeley, and S. Zuyev, "Filters for spatial point processes," Cambridge University Engineering Department, Tech. Rep. CUED/F-INFENG/TR-591, 2007.