Ground Multiple Target Tracking with a Network of Acoustic Sensor Arrays Using PHD and CPHD Filters

Emre Özkan, Mehmet B. Guldogan, Umut Orguner, Fredrik Gustafsson
Division of Automatic Control
Department of Electrical Engineering
Linköping University, Sweden
Email: {emre, bguldogan, umut, fredrik}@isy.liu.se

Abstract—In this paper, we explore the potential of networked microphone arrays for multiple target tracking. Tracking is accomplished by using the direction-of-arrival (DOA) estimates of multiple microphone arrays. Each microphone array obtains the DOA estimates by using the wideband extensions of the multiple signal classification (MUSIC) technique. Based on these DOA estimates, multi target tracking is done by using Probability hypothesis density (PHD) and cardinalized probability hypothesis density (CPHD) algorithms. The results show that the CPHD performs better than the PHD on real data as it estimates the number of targets more accurately.

Keywords: Acoustic ground multitarget tracking, bearing estimation, PHD filter, CPHD filter.

I. INTRODUCTION

Localization and tracking of multiple ground targets, e.g., cars and trucks, using unattended acoustic sensor arrays, where sensors are placed in an array-of-arrays configuration, is a very challenging problem due to several reasons. The difficulty arises from many sources such as the nonstationary and wideband structure of target acoustic signatures, wind effects, scattering by turbulence, severe attenuation, near-field effects and acoustic clutter [4]. Moreover, DOA estimation, data association and tracking problems become further complicated in the case of closely spaced ground targets.

Here, we are concerned with the tracking of multiple ground targets using an acoustic unattended ground sensor network. Each acoustic array has local processing capability and provides bearing measurements. The acoustic spectra of ground vehicles are nonstationary and wideband. There exist harmonics in the spectra. For proper acoustic multi target tracking, it is important to get benefit of the multi spectral content from the acoustic targets and determine frequency peaks to improve the accuracy and consistency of the angle estimates. In the literature, there exist different wideband DOA estimation techniques to detect moving ground targets from their acoustic spectra [1], [3], [5], [6], [7]. Wideband version of the classical subspace based multiple signal classification (MUSIC) technique is one of the good techniques that have provided successful results in resolving multiple moving targets. Conventional delay-sum beamformers and narrowband version of the MUSIC technique have poor performances in wideband multi-target scenarios [7]. Moreover, in [2], performance of different wideband MUSIC approaches on real acoustic measurements are analyzed.

Among possible methods proposed for the solution of multi target tracking problem, Mahler’s probability hypothesis density (PHD) filters have gained popularity in the field [13]. The PHD filter is a set theory based method which approximates the multi-target posterior distribution by propagating its first order moment. The PHD formulations avoid the data association problem and do not involve keeping the identity of the targets. The Cardinalized PHD is later proposed in order to jointly propagate the posterior distribution of the number of targets (cardinality distribution) and the posterior PHD [12] [14]. Its main advantage over the PHD filter is its ability to produce more stable and accurate estimate of the number of targets than the PHD filters which is preferable in heavy clutter scenarios. CPHD filters are also claimed to be more robust in case of missed detections [16].

The PHD filter and its variants have been applied to many applications [8], [9], [10], [11], [19]. To the best of our knowledge, our work is the first attempt where CPHD and PHD filters are used in multiple ground vehicle tracking using acoustic sensor arrays. In this paper, we tested the performances of both the PHD and the CPHD filters on a real multiple ground target scenario using unattended acoustic sensor arrays.

The paper is organized as follows. In the Section-II, signal model is provided. In the Section-III, wideband DOA estimation is discussed and details of the used techniques are presented. In the Section-IV, multi sensor multi target problem is formulated and PHD and CPHD filter recursions are reviewed. In the Section-V, some implementation details are pointed out. Finally, experimental setup and some results on multi target tracking are presented in the Section-VI.

II. DATA MODEL

Consider that, \( N_T \) wideband sources impinge on a set of acoustic sensors and the sensors are placed in an array-of-arrays configuration containing several small aperture arrays distributed over an area. There exists \( N_A \) number of arrays in
the area and each array, \( k = 1, \ldots, N_A \), contains \( N_S \) number of omnidirectional sensors. Sensors, \( m = 1, \ldots, N_S \), are uniformly oriented on a circle of radius \( r_k \). The sensor positions are \( \mathbf{r}_{k,m} = r_k [\cos(m \Delta), \sin(m \Delta)]^T \) where \( \Delta = 2\pi/N_S \) and \( T \) denotes the transpose. We assume that the arrays and the sources are coplanar. The received signal by the \( m \)th sensor of the \( k \)th array can be expressed as:

\[
    z_{k,m}(t) = \sum_{n=1}^{N_T} \zeta_n s_n(t - \mathbf{r}_{k,m}^T \mathbf{a}_{n,k}) + v_{k,m}(t), \tag{1}
\]

where \( \zeta_n \) is the complex scaling factor of the \( n \)th target containing all the attenuation and phase terms, \( s_n(t) \) represents the waveform of the \( n \)th target, \( \mathbf{a}_{n,k} \equiv (1/c)[\cos(\theta_{n,k}), \sin(\theta_{n,k})]^T \) is the slowness vector of the \( n \)th target (\( c \) is the speed of sound, \( \theta_{n,k} \) is the DOA of the \( n \)th target onto the \( k \)th array) and \( v_{k,m} \) is the complex additive noise, which is assumed to be spatially and temporally white Gaussian distributed with covariance \( \sigma^2 \). In the following formulations, since \( \zeta_n \) has no effect on DOA estimation, it is set to \( \zeta_n = 1 \).

In order to process the nonstationary acoustic signals, the short-time Fourier transform (STFT) is used. The STFT of a signal \( s(t) \) is defined as:

\[
    s(f, t) = \sum_{\tau=1}^{T} s(t - \tau) w(\tau) e^{-j2\pi f \tau}, \tag{2}
\]

where \( w(\tau) \) is a window function of length \( T \). Window size should be chosen such that the practical stationarity of the signal is preserved. With this definition, STFT of the sensor output \( z_{k,m}(t) \) is

\[
    z_{k,m}(f, t) = \sum_{n=1}^{N_T} s_n(f, t) e^{-j2\pi f \mathbf{r}_{k,m}^T \mathbf{a}_{n,k}} + v_{k,m}(f, t). \tag{3}
\]

If each sensor output in the \( k \)th array are combined in a vector, we can express the STFT of the \( k \)th array as

\[
    \mathbf{z}_k(f, t) = \sum_{n=1}^{N_T} s_n(f, t) \mathbf{a}_k(f, \theta_n) + \mathbf{v}_k(f, t). \tag{4}
\]

where \( \mathbf{a}_k(f, \theta_n) \) is the unit steering vector of the \( k \)th array and is defined as

\[
    \mathbf{a}_k(f, \theta_n) = [\ e^{-j2\pi f \mathbf{r}_{k1}^T \mathbf{a}_{n,k}}, \ldots, e^{-j2\pi f \mathbf{r}_{kN_S}^T \mathbf{a}_{n,k}} \ ]^T. \tag{5}
\]

More compact matrix form of the equation (4) is

\[
    \mathbf{z}_k(f, t) = \mathbf{A}_k(f, \theta) \mathbf{s}(f, t) + \mathbf{v}_k(f, t), \quad k = 1, \ldots, N_S \tag{6}
\]

where \( \mathbf{A}_k(f, \theta) = [ \mathbf{a}_k(f, \theta_1), \ldots, \mathbf{a}_k(f, \theta_{N_T})] \) is the \( N_S \times N_T \) steering matrix of the \( k \)th array and \( \mathbf{s}(f, t) = [s_1(f, t), \ldots, s_{N_T}(f, t)]^T \) is the vector of targets in the frequency domain. In time frequency domain, (6) tells us that the nonstationary and wideband target DOA estimation is similar to the stationary and narrowband case at each time \( t \) and at each frequency \( f \).

Power spectral density matrix for the \( k \)th sensor array output can be written as

\[
    \mathbf{R}_{z_k}(f, t) = \mathbf{E} [\mathbf{z}_k(f, t) \mathbf{z}_k(f, t)^H] = \mathbf{A}_k(f, \theta) \mathbf{R}_s(f) \mathbf{A}_k(f, \theta)^H + \mathbf{R}_v(f, t). \tag{7}
\]

where \( \mathbf{E}[\cdot] \) denotes the expectation and the superscript \( H \) denotes the Hermitian transpose operator. \( N_T \times N_T \) matrix \( \mathbf{R}_s(f) = \mathbf{E} [\mathbf{s}(f) \mathbf{s}(f)^H] \) and the \( N_S \times N_S \) matrix \( \mathbf{R}_v(f) = \mathbf{E} [\mathbf{v}(f) \mathbf{v}(f)^H] \) are the power spectrum matrices of targets and the noises, respectively.

### III. WIDEBAND BEARING ESTIMATION

In this section, we present two different MUSIC based DOA estimation techniques called as coherent and incoherent wideband MUSIC techniques [1], [2]. As mentioned in the previous section, the data was segmented into fixed blocks before the processing step and stationarity is assumed over each block. Therefore, in the following equations, \( \mathbf{R}_{z_k}(f, t) \equiv \mathbf{R}_{z_k}(f) \) is used. With this assumption, (7) is rewritten again:

\[
    \mathbf{R}_{z_k}(f) = \mathbf{A}_k(f, \theta) \mathbf{R}_s(f) \mathbf{A}_k(f, \theta)^H + \mathbf{R}_{v}(f). \tag{8}
\]

Sensor array outputs are decomposed into narrowband components by adaptively selecting the \( N_T \) frequency components for wideband processing. Frequency selection can be performed either by simple thresholding based on frequency bin signal-to-noise ratio (SNR) or by using more complex approaches. For each frequency bin, \( \ell = 1, \ldots, N_T \), power spectrum matrix is formed as:

\[
    \mathbf{R}_{z_k}(f_\ell) = \mathbf{A}_k(f_\ell, \theta) \mathbf{R}_s(f_\ell) \mathbf{A}_k(f_\ell, \theta)^H + \mathbf{R}_{v}(f_\ell). \tag{9}
\]

Eigenvalue decomposition (EVD) of \( \mathbf{R}_{z_k}(f_\ell) \) is given as

\[
    \mathbf{R}_{z_k}(f_\ell) = \mathbf{U}(f_\ell) \Lambda(f_\ell) \mathbf{U}(f_\ell)^H(f_\ell) \tag{10}
\]

where \( \Lambda(f_\ell) = \text{diag} \{ \lambda_1(f_\ell), \ldots, \lambda_{N_S}(f_\ell) \} \) is a diagonal matrix with \( \lambda_1(f_\ell) \geq \cdots \geq \lambda_{N_S}(f_\ell) \geq 0 \), being nonnegative eigenvalues, and \( \mathbf{U}(f_\ell) \) is a unitary matrix \( \mathbf{U}(f_\ell) = [u_1(f_\ell), u_2(f_\ell), \ldots, u_{N_S}(f_\ell)] \) containing the corresponding \( N_S \) eigenvectors. \( \mathbf{U}_s(f_\ell) \) contains the signal eigenvectors, and \( \mathbf{U}_n(f_\ell) \) contains the noise eigenvectors with \( \Lambda_n(f_\ell) = \sigma^2 I \) corresponding to the \( N_S - 1 \) smallest eigenvalues. In this paper, maximum eigenvalue, \( \Lambda(f_{11}) \), is selected to constitute a signal subspace. In other words, a single frequency, \( f_1 \), is assumed to be occupied by a single target only. This assumption has a practical meaning, since different wideband targets are not likely to occupy all of the same frequency bins in any processing time interval and the frequency bins change as time progresses [2]. For both approaches, incoherent and coherent, MUSIC spectrum is used to compute the spatial beam pattern. The main steps of the DOA estimation are as follows:

- First, adaptively select the narrowband frequency bins on windowed stationary data.
- Apply the incoherent or the coherent processing with MUSIC.
Estimate the DOAs of each target from the resultant spatial spectrum.

In incoherent processing, spatial spectrum is obtained by adding each spatial spectrum for each frequency bin, \( \ell = 1, \ldots, N_F \):

\[
P_{k}^{nc}(\theta) = \sum_{\ell=1}^{N_F} P_{k}(f_{\ell}, \theta)\tag{12}
\]

\[
= \sum_{\ell=1}^{N_F} a_{k}(f_{\ell}, \theta)U_{n}(f_{\ell})U_{n}(f_{\ell})^{H}a_{k}(f_{\ell}, \theta)
\]

In coherent processing, the power spectrum matrix at each adaptively selected frequency bin is steered to a certain look angle by a diagonal steering matrix and subspace fitting is conducted on the focused new covariance matrix unlike the incoherent processing. This approach is also known as the steered covariance method \([1]\). The obtained steered covariance can be written as

\[
R_{a_{k}}(\theta) = \sum_{\ell=1}^{N_F} \Phi(f_{\ell}, \theta)R_{a_{k}}(f_{\ell})\Phi(f_{\ell}, \theta)^{H},\tag{13}
\]

where \( \Phi(f_{\ell}, \theta) = \text{diag} \{ a_{k}(f_{\ell}, \theta) \} \) is \( N_{S} \times N_{S} \) diagonal matrix with elements of \((5)\). Similar to \((11)\), EVD of \( R_{a_{k}} \) is calculated:

\[
R_{a_{k}}(\theta) = U_{s}(\theta)\Lambda_{s}(\theta)U_{s}(\theta)^{H} + U_{n}(\theta)\Lambda_{n}(\theta)U_{n}(\theta)^{H}\tag{14}
\]

and it is assumed that only one target exists for each angle \( \theta \). Therefore, a signal subspace is formed from one eigenvector corresponding to the largest eigenvalue and noise subspace is formed from the remaining \( N_{S} - 1 \) eigenvectors. The coherent wideband MUSIC angle spectrum is given by

\[
P_{k}(\theta) = \frac{1}{a_{k}(f_{0}, \theta)U_{n}(\theta)U_{n}(\theta)^{H}a_{k}(f_{0}, \theta)}\tag{15}
\]

where \( U_{n}(\theta) \) is unitary noise subspace and \( f_{0} \) is the center frequency of frequency band that is under consideration. After computing the angle spectrum, the DOA estimates are found from the peaks of the spectrum.

IV. Multi Target Tracking

Once the DOA estimates are obtained from the acoustic data, suitable tracking scheme should be constructed in order to track the targets of interest. In this section we will consider the problem of tracking an unknown number of ground targets by using DOA estimates collected from multiple sensors.

Our basic assumptions state that each target state evolves according to a known Markov transition density \( p(x_{t}|x_{t-1}) \) and will generate measurements at each sensor (if detected) according to a known measurement equation inducing the single measurement likelihood \( g(y_{t}|x_{t}) \). For a given road map we will define on-road position and on-road velocity of each target in order to reduce the dimensionality of the problem. On-road state vector of target \( j \) at time \( t \) is defined as,

\[
x_{j}^{t} = [p_{j}^{t} \quad v_{j}^{t}]^{T},\tag{16}
\]

where \( p_{j}^{t} \) is the position and \( v_{j}^{t} \) is the velocity of the target \( j \). Global state vector of the target \( j \) at time \( t \) is defined as,

\[
x_{global,t}^{j} = [p_{x,t}^{j} \quad p_{y,t}^{j} \quad v_{x,t}^{j} \quad v_{y,t}^{j}]^{T}.\tag{17}
\]

We assume that there exist a suitable coordinate transformation \( \Xi \) which maps on-road state vectors to the global state coordinates.

\[
x_{global,t}^{j} = \Xi(x_{j}^{t})\tag{18}
\]

The state space as defined on the road map is then quantized on grids where we will propagate the PHD and CPHD by using point mass filters. The target motion on the road is modeled by a constant velocity model.

\[
x_{j}^{t} = \begin{pmatrix} 1 & T \\ 0 & 1 \end{pmatrix} x_{j}^{t-1} + \begin{pmatrix} T^{2}/2 \\ T \end{pmatrix} e_{t}\tag{19}
\]

where \( e_{t} \sim \mathcal{N}(0, Q) \) is Gaussian noise term accounting for possible accelerations. Measurements are the noisy DOA estimates of the sensors.

\[
y_{t} = \arctan \left( \frac{p_{y,t} - p_{y,sensor}^{t}}{p_{x,t} - p_{x,sensor}^{t}} \right) + \epsilon_{\theta,t}\tag{20}
\]

where \( p_{x,t}, p_{y,t} \) indicates the position of the target, \( p_{x,sensor}^{t}, p_{y,sensor}^{t} \) indicate the position of the sensor in global coordinate system. \( \epsilon_{\theta,t} \) is the measurement noise which is assumed to be Gaussian, i.e., \( \epsilon_{\theta,t} \sim \mathcal{N}(0, R) \). Having defined the basics of the multitarget problem, we are going to review the PHD and CPHD filters in the following subsection.

A. PHD and CPHD recursions

The PHD filter is a finite set statistics (FISST) based approach where the single target Bayesian recursions are extended to multitarget states updated by multitarget measurements. The multitarget state is modeled by a random finite set (RFS). The posterior distribution of the multitarget state RFS can be computed by Bayesian recursion. Because of the combinatorial nature of the multitarget tracking problem, it is computationally expensive to propagate the full posterior. Therefore its first order statistical moment is propagated instead. The PHD recursions \([13]\) are computed in two steps, namely the prediction and the correction steps. The PHD recursions can be written as follows.

**Prediction:**

\[
D_{t|t-1}(x) = b(x) + \int f(x|\hat{x})p_{s}(x)D_{t-1|t-1}(\hat{x})d\hat{x}\tag{21}
\]

**Correction:**

\[
D_{t|t}(x) = (1 - p_{D}(x))D_{t|t-1}(x) + p_{D}(x)\sum_{j=1}^{m_{t}} \frac{g(y_{t}^{j}|x)}{c(y_{t}^{j})} + \int p_{D}(x)g(y_{t}^{j}|x)D_{t|t-1}(x)d\hat{x}D_{t|t-1}(x)\tag{22}
\]

where \( D_{t|t-1}(x) \) and \( D_{t|t}(x) \) are the predicted and updated PHD’s at time \( t \), \( Y_{t} = y_{t}^{1}, \ldots, y_{t}^{m_{t}} \) is the set of measurements from a single sensor at time \( t \), \( m_{t} \) is the number of measurements, \( f(x|\hat{x}) \) is the single target state transition density,
$g(y|x)$ is the measurement likelihood, $p_s(.)$ is probability of survival of a target, $b(.)$ is the spontaneous target birth intensity, $c(.)$ is the clutter intensity, $p_D(.)$ is the probability of detection.

CPHD filter propagates the probability distribution of number of targets (cardinality) jointly with the PHD. CPHD recursions are more complicated than that of PHD and include additional update steps for the cardinality distribution.

**Prediction:**

$$D_{t|t-1}(x) = b(x) + \int f(x|x')p_sD_{t-1|t-1}(x)dx$$  \hspace{1cm} (23)

$$p_{t|t-1}(n) = \sum_{\hat{n}} p_{t-1|t-1}(\hat{n}) M(n, \hat{n})$$  \hspace{1cm} (24)

$$M(n, \hat{n}) = \sum_{i=0}^{\min(n, \hat{n})} p_{birth}(n-i) \left( \frac{\hat{n}}{i} \right) (1-p_s)^{\hat{n}-i}p_s^i$$  \hspace{1cm} (25)

where $p_{t|t-1}(\cdot)$ is the cardinality distribution of the targets at time $t$, $p_{birth}$ is the target birth cardinality distribution and $p_s$ is the probability of survival of a target which we assumed state independent. At any time $t$, expected number of targets can be computed from the cardinality distribution by taking the expectation.

$$n_{t|t} = \sum_{n=0}^{\infty} np_{t|t-1}(n)$$  \hspace{1cm} (26)

The correction step of the CPHD filter can be written in terms of relevant likelihood ratios as follows [15].

**Correction:**

$$D_{t|t}(x) = \left[ (1-p_D(x)) \frac{L(Y_t|D^-)}{L(Y_t)} + \frac{L(Y_t|D)}{L(Y_t)} \right] D_{t|t-1}(x)$$  \hspace{1cm} (27)

$$p_{t|t}(n) = \frac{L(Y_t|n)}{L(Y_t)} p_{t|t-1}(n)$$  \hspace{1cm} (28)

where $D$ and $D^-$ denote the conditions target detected and not detected respectively. The likelihood ratios are computed as,

$$\mathcal{L}(Y_t|D^-) = \frac{1}{n_{t|t-1}} \sum_{j=0}^{m_t} \alpha_t^{(j+1)} \beta_t^{(j)} \sigma_j(L_t^{(1)}, \ldots, L_t^{(m_t)})$$  \hspace{1cm} (29a)

$$\mathcal{L}(Y_t|D) = \frac{1}{n_{t|t-1}} \sum_{s=1}^{m_t} \alpha_t^{(j)} \beta_t^{(j)} \sigma_j(L_t^{(1)}, \ldots, L_t^{(m_t)}) \ln \mathcal{L}(y_t^{(s)}|x)$$

$$\sum_{j=1}^{m_t} \alpha_t^{(j)} \beta_t^{(j)} \sigma_j(L_t^{(1)}, \ldots, L_t^{(m_t)}) \sigma_j(L_{t-1}^{(1)}, \ldots, L_{t-1}^{(m_t)})$$  \hspace{1cm} (29b)

$$\mathcal{L}(Y_t|n) = \sum_{j=0}^{\min(m_t, n)} \beta_t^{(j)} \left( \frac{n!}{(n-j)!} (1-p_D)^{n-j} \right)$$  \hspace{1cm} (29c)

where

$$\alpha_t^{(j)} = \frac{\sum_{n=j}^{\infty} n! p_{t|t-1}(n)(1-p_D)^{n-j}}{n_{t|t-1}}$$  \hspace{1cm} (30)

$$\beta_t^{(j)} = p_c(m_t-j) (m_t-j)! \lambda^{-j}$$  \hspace{1cm} (31)

$$n_{t|t-1} = \sum_{n=0}^{\infty} np_{t|t-1}(n)$$  \hspace{1cm} (32)

$p_c(m)$ is the cardinality distribution of false alarms, $\lambda$ is the clutter density. The likelihood of target originated measurements are defined as,

$$L_t^{(s)} = \frac{1}{n_{t|t-1}} \int g(y_t^{(s)}|x)p_D(x)D_{t|t-1}(x)dx$$  \hspace{1cm} (33)

The function $\sigma_j(A)$ is the elementary symmetric function which is defined as the sum of all possible products of elements of the set $A$ with $j$ different factors.

$$\sigma_j\{a_1, \ldots, a_m\} = \sum_{1 \leq i_1 < \cdots < i_j \leq m} a_{i_1}a_{i_2}\cdots a_{i_j}$$  \hspace{1cm} (34)

and $\sigma_0 = 1$. If the probability of detection $p_D$ is state dependent, then the average value for $p_D$ is substituted in equations (29d) and (30).

$$\tilde{p}_D = \frac{1}{n_{t|t-1}} \int p_D(x)D_{t|t-1}(x)dx$$  \hspace{1cm} (35)

**V. IMPLEMENTATION**

It is possible to implement PHD filters with different methods. SMC implementation of the PHD is done in [21]. Gaussian approximation of PHD and CPHD filters are also possible [15], [20], [22]. Here we will make the inference by using point mass filters where we divide the on-road state space of the targets into grids and propagate PHD and CPHD. At a given time $t$, PHD will be approximated on this grid with a set of $N$ particles and their weights.

$$D_{t|t} = \sum_{i=1}^{N} w_{t|i} \delta(x - x_t^i)$$  \hspace{1cm} (36)

The prediction and correction steps will be applied to particle weights sequentially. Before presenting the pseudo code of the algorithms, we describe the common stages of PHD and CPHD updates. Assume we have the approximation of PHD from the previous time step $t-1$.

$$D_{t-1|t-1} = \sum_{i=1}^{N} w_{t-1|i-1} \delta(x - x_t^i)$$  \hspace{1cm} (37)

According to the equations (21) and (23) the predicted PHD will be

$$D_{t|t-1} = \sum_{i=1}^{N} w_{t|i-1} \delta(x - x_t^i)$$  \hspace{1cm} (38)

$$w_{t|i-1} = w_{birth} + \sum_{j=1}^{N} p_s(x_t^j)f(x_t^j|x_t^{i-1})w_{t-1|i-1}$$  \hspace{1cm} (39)
For the point mass implementation of the PHD \( f(x^j_t| x^i_{t-1}) \) terms are fixed and can be computed offline. The prediction update of the cardinality distribution of CPHD is done using the equations (24) and (25). Here the maximum number of targets expected to be observed in the field of interest is set to a limit, in other words the cardinality distribution is truncated to zero after a finite value such that the computations are tractable. For constant target birth and target survival probabilities \( M \) matrix can be computed a priori.

A. Multiple sensor PHD

For the sake of clarity, so far we have considered the single sensor update of the PHD filter. Extending the PHD filters for multisensor multi target tracking is still an open problem in the literature. In [17] and [18] Mahler discusses possible ways of extending PHD algorithms for multiple sensors. Here we will use iterative corrector approach, in which the sensors apply the correction step over the corrected PHD of the previous sensor.

B. Target extraction

Target extraction is done from the estimated PHD by first estimating the number of targets, then detecting the peaks. In the PHD filter, the expected number of targets can be found in a specific region of the state space by taking the integral of the PHD over that region. The expected number of targets on the road is equal to the sum of the weights.

\[
\tilde{N}_T = \sum_{i=1}^{N} w^{(i)}_t
\]

Estimated number of target can be chosen as the closest integer to that sum. In CPHD filter, the sum of the weights is equal to the mean of the cardinality distribution, which is the expected number of targets. Furthermore MAP estimate of the number of targets can be obtained from the maxima of the cardinality distribution. Given the estimated number of targets, as many peaks are detected from the estimated PHD by using a simple peak detector. The detected peaks correspond to the estimated target states.

VI. EXPERIMENTAL RESULTS

In the experiments, real acoustic network data is collected in an area close to town Skövde, Sweden. The map of the area is shown in Figure 1 along with the road segment information and microphone array positions. The road segment information is composed of connected linear smaller road segments. The on-road position coordinates are marked in the figure at each 50 meters. In total, \( N_A = 4 \) acoustic microphone arrays are deployed, each has \( N_S = 3 \) uniformly spaced microphones with a separation of \( 2\pi/3 \) on a circle of radius 0.15 meters. Two microphone arrays have the same location such that the one is located on the ground level and the other one is 1.4 meters above the ground level. Acoustic microphone array locations on the map are denoted with \( \times \) and \( \circ \) in Figure 2. The acoustic spectra of moving ground vehicles are wideband and nonstationary. There exist harmonics in the spectra and most of the detectable spectral peaks are below 200 Hz. Since the acoustic signatures are corrupted due to wind noise, spectral peaks below 30 Hz are not considered. Data collecting rate of each microphone in an array is 48 kHz. DOA’s are estimated by using two methods, coherent and incoherent wideband MUSIC techniques. Tracking results by using different DOA estimates are given. Three vehicles, each moving with a constant velocity of 30 km/h are used in the experiment. In the first 50 seconds of the experiment, one vehicle is moving on the road between two ends of the road. In the remaining 40 seconds two vehicles are moving. Reference positions of the targets are measured with differential GPS (DGPS) with sampling rate of 1 Hz. The road grid resolution is taken as 2m. The total arclength of the road is 708m. Velocity grid resolution is 1m/s between the values -10 m/s and 10 m/s. The common parameters of PHD and CPHD filters are chosen to be the same. \( p_b = 0.99, \lambda = 0.1 rad^{-1} \). Probability of detection is state dependent with maximum value 0.8 and weighted according to the road map and sensor positions. Maximum number of targets for the cardinality distribution of CPHD is \( T_{MAX} = 20 \). Clutter cardinality is Poisson with mean value 1, and \( p_{birth} = 0.001 \). Measurement standard deviation is taken as 0.2 rad. In Figures 3 and 4, the estimated number of targets are depicted for PHD and CPHD using incoherent wideband MUSIC DOA estimates. As can be observed from the figures, CPHD algorithm is better in estimating the number of targets than PHD. The variance of the estimate of number of targets of the PHD algorithm is higher than the CPHD algorithm as expected. In Figures 5 - 6 the estimated target trajectories are depicted. At the beginning of the scenario, there exists only one target and the CPHD algorithm declares that there is a target earlier than the PHD. CPHD also preserves the track continuity better for this target. In the second part of the scenario there are two crossing targets. After the crossing takes place, the detected DOA estimates of one target fall far away from the true target state. PHD drops the target at this point.
Table II

PSEUDO CODE FOR CPHD

Multi Sensor CPHD Filter:
  * **Initialization:**
    - For each particle \( i = 1, \ldots, N \) do
      - Set Maximum number of targets, \( T_{\text{max}} \).
      - Set initial weights \( \bar{w}^{(0)}_i \).
      - Set initial cardinality distribution
        \( p_0(n), n = 1 : T_{\text{max}} \).
      - Compute \( M_{\text{diag}}(\cdot, \cdot) \) matrix according to equation (25)
  * **Iterations:**
    - For \( t = 1, 2, \ldots \) do
      - **Prediction Step:**
        - For each particle \( i = 1, \ldots, N \)
          * Compute predicted weights:
            \[ \bar{w}_i^t = w_{\text{birth}} + \sum_{j=1}^{N} p_s(x_j^t) f(x_j^t|x_{i}^{t-1}) w_{i}^{t-1}(t-1) \]
          - Compute predicted cardinality
            \[ \bar{p}_i^t(n) = \sum_{j=1}^{T_{\text{max}}} p_s(x_j^t) f(x_j^t|x_{i}^{t-1}) w_{i}^{t-1}(t-1) \]
        - For each sensor \( q = 1, \ldots, N_s \) do
          * For each measurements of the sensor \( j = 1, \ldots, m_q^n \)
            - Compute \( L_j^t(\cdot) = \sum_{x=1}^{N} \hat{p}(y|\cdot) f(x_j^t|x_{i}^{t-1}) \bar{w}_i^t \)
          - Compute \( L(Y_1|D^-), L(Y_2|D), L(Y_2|n), L(Y_1) \) according to equations (29a-d)
        - Update weights:
          \[ w_{i}^t = \begin{cases} 
            1 - p_D & L(Y_1|D^-) + L(Y_2|D) \\
            L(Y_2|n) & L(Y_1) 
          \end{cases} \]
      - **Correction Step:**
        - For the correction of the next sensor,
          * set \( \bar{w}_i^t = w_{i}^t \)
          * set \( p_{i|t-1}^t(n) = \bar{p}_i^t(n) \)
          * set \( n_{i|t-1}^t(n) = n_{i|t-1}^t(n) \)

Table III

PSEUDO CODE FOR CPHD

whereas CPHD still keeps the track and correctly estimates the cardinality. Apart from these points, the estimated target trajectories for the both algorithms are satisfactory. The DOA estimates of this scenario are shown in Figure 7. In Figures 8 - 11, the results for the use of coherent wideband MUSIC technique are given. The DOA estimates of coherent MUSIC method are obtained from relatively flat spectrum therefore less number of peaks are detected as false alarms. The sensors can not detect the vehicles at all scans and that would cause track discontinuities. This tendency can be observed in 8 - 9 where CPHD’s estimates are more stable. On the other hand, the estimated state trajectories are satisfactory for both algorithms and they both manage to track the crossing targets successfully. The clutter density for PHD and CPHD filters is set to \( \lambda = 0.05 \text{rad}^{-1} \) for the coherent method due to less number of false alarms and the other parameters are kept the same.

VII. CONCLUSIONS

In this work, we considered the problem of multiple sensor tracking of multiple ground vehicles using acoustic sensor arrays. The performance of the PHD and CPHD filters are tested on real data. Advantages of using CPHD filters instead of PHD filters have been discussed in the literature. Our work here stands as an example of illustrating the advantages of using relatively more complicated and computationally demanding CPHD filters over the PHD filters based on real data experiments. The detection of multiple targets by acoustic sensor arrays is a difficult task such that one target close to the array can conceal the other targets. In such cases, one should take this occlusion in the acoustic domain into account and utilize better state dependent probability of detections, for which further research is necessary. It is also worth mentioning that both the PHD and CPHD algorithms have their own characteristic problems for missed detections (see: e.g. [23]). The effect of these errors in the case of multiple sensors is...
also a topic of interest.

ACKNOWLEDGMENTS

The authors gratefully acknowledge fundings from the Swedish Research Council VR in the Linnaeus Center CADICS, and Swedish Foundation for Strategic Research in the center MOVIII. The strategic motivation and practical relevance stem from the FOI Center for Advanced Sensors, Multisensors and Sensor Networks (FOCUS) funded by the Swedish Governmental Agency for Innovation Systems (VINFÖR), and the Knowledge Foundation (KK-stiftelsen). The authors would also like to specifically thank Fredrik Gunnarson at LiU, David Lindgren and Hans Habberstad at FOI for contributing with the dataset.

REFERENCES


Figure 7. DOA estimates of incoherent wideband MUSIC (Four sensors)

Figure 8. Estimated number of Targets by PHD.

Figure 9. Estimated number of Targets by CPHD.

Figure 10. Estimated Target trajectories of PHD.

Figure 11. Estimated Target trajectories of CPHD.


