Histogram PMHT with Target Extent Estimates Based on Random Matrices

Monika Wieneke
Department Sensor Data and Information Fusion
Fraunhofer FKIE / University of Bonn
D-53343 Wachtberg, Germany
Email: monika.wieneke@fkie.fraunhofer.de

Samuel J. Davey
Intelligence Surveillance and Reconnaissance Division
Defence Science and Technology Organisation
Edinburgh, SA 5111, Australia
Email: samuel.davey@DSTO.Defence.gov.au

Abstract—Conventional tracking approaches are based on the assumption that the targets to be tracked are point targets and that the measurements to be processed are point measurements. However, when a sensor provides image data of high resolution in which targets might be distributed over several display cells, neither assumption is suitable. In such applications the estimation of the target extent and the utilization of the complete image frame are crucial to achieve good tracking performance.

Recently, a Bayesian filter for single extended object tracking based on random matrices has been proposed. In this approach ellipsoidal object extents are modeled by random matrices and treated as additional state variables. This article deals with an integration of random matrices into the Histogram Probabilistic Multi-Hypothesis Tracker. The novel approach tracks multiple extended targets directly in an image sequence without previous point measurement extraction. The superiority of the algorithm is proven by simulations.

Keywords: Extended Object Tracking, Random Matrices, Histogram PMHT, Track-Before-Detect, Data Assignment, Multi-Target Tracking, Intensity Tracking, Estimation.

I. INTRODUCTION

In realistic tracking scenarios the targets of interest are not usually single points in space but rather extended objects with differing and time-varying shapes. The demand for techniques with the ability to deal with extended objects arises in the following situations: First, modern sensor technology makes it possible to deploy high-resolution sensors to gather data from the objects. For such sensors, the size of the data cells within a sensor frame may be small compared with the size of the target, leading to a response over multiple cells. This is naturally emphasized the closer a target is positioned in front of the sensor. Second, there may be targets that move collectively in a dense group formation. Such constellations make it rather difficult to resolve all data assignment conflicts between the individual group members. In this case it is more efficient to treat each group as a single extended object to be tracked without estimating the exact state of each group member.

An extended object typically comprises multiple scattering centers that give rise to a relatively large and often fluctuating number of measurements. To produce stable tracks, a tracking algorithm has to take into consideration that the collection of scattering centers that are responsible for the measurements might change from scan to scan. Such a property is naturally not provided by the traditional tracking algorithms, since they are all based on the assumption that the targets to be tracked are point sources. Therefore there is an increasing need for algorithms that have the ability to recognize extended objects as individual units, as well as the ability to initiate, to maintain and finally to delete extended object tracks.

The nature of extended objects is also ill suited to the conventional point measurement extraction methods. A common technique is to apply a detection algorithm to extract potential target points from the current sensor frame and then send the points to the tracker module. The detection algorithm typically applies a threshold to the data to treat those cells that exceed the threshold as point measurements, perhaps with additional interpolation methods to improve accuracy. This is acceptable if the signal-to-noise ratio (SNR) is high. For low SNR targets the threshold has to be decreased to allow sufficient probability of target detection. However, a low threshold also gives a high rate of false detections, which causes the tracker to form false tracks. For extended objects, a point measurement detector may also produce a cluter of measurements from a single object or measurements which fluctuate from one part of the object to another.

An alternative approach is represented by a paradigm called Track-Before-Detect (TkBD). TkBD combines target detection and estimation by removing the detection algorithm from the process and supplying the whole sensor frame directly to the tracker. This improves track accuracy and allows the tracker to follow low SNR targets. The Histogram Probabilistic Multi-Hypothesis Tracker (H-PMHT) [2]–[4] proposed by Streit is a parametric TkBD algorithm, which means that it assumes the superposition of the power from scattering sources and probabilistically associates the received power in each sensor cell with target and clutter models. After the association phase it can represent the contribution due to a particular component as an equivalent point measurement and then exploit algorithms for point measurement tracking. H-PMHT is naturally a multi-target algorithm.

In this paper we extend the H-PMHT algorithm by introducing random matrices to jointly estimate both the extent matrix and the kinematical state vector of multiple ellipsoidal targets. The application of random matrices for recursively estimating the joint state of a single extended object was first presented by
Koch [1]. Here we show that the H-PMHT with incorporated extent estimates is a natural multi-target extension of the original Bayesian approach in [1]. After the data association phase the novel H-PMHT exploits the single extended object filter to derive improved estimates for the kinematical states and extents of each target. Moreover, the derived approach is the first TkBD algorithm for extended object tracks based on random matrices. It is to extract, maintain and delete the extended object tracks, and to ignore clutter measurements. We show the superiority of our approach by means of simulations.

II. RELATED WORK

In [1] a Bayesian filter for extended object tracking based on random matrices is proposed. Within this approach, the physical extent of an object is modeled by an ellipse whose parameters are a random matrix and are treated as additional state variables to be estimated. However, the approach allows for only a single target. The first multi-target version of the filter in [1] was presented by Wieneke and Koch. Here we show that the H-PMHT with incorporated data association phase, the PMHT with extent estimates runs a bank of extended object trackers. The filter in [1] was proven to be a special case of the PMHT in [6].

III. BAYESIAN SINGLE EXTENDED TARGET FILTER

To integrate the target extent as an additional estimate into the Bayesian tracking recursion, the kinematical state $x_t$ at each time $\tau_t$ has to be augmented by a symmetric positive definite (SPD) random matrix $X_t$ whose estimates are derived from the measurements.

The kinematical state variable $x_t$ of the target centroid at time $\tau_t$ is, given by $x_t = \begin{bmatrix} r_t^T, \xi_t^T \end{bmatrix}^T$ consisting of spatial state components $r_t \in \mathbb{R}^d$. The component $r_t$ is the corresponding velocity. The SPD matrix $X_t \in \mathbb{R}^{d \times d}$ designates the current ellipsoidal object extent. The dimension of the kinematical state vector $x_t$ is therefore $s \cdot d$, with $s = 1$ describing up to which derivative the object kinematics is modeled. In this work we estimate position and velocity, which means $s = 2$. Let $Z_t = [z_t^1, \ldots, z_t^{N_t}]$ be the vector of $N_t$ measurements at time $\tau_t$, where each measurement $z_t^i \in \mathbb{R}^d$ is assumed to be a Cartesian position. $Z^t$ denotes all measurement vectors from scan 1 up to the current scan $t$. Furthermore, since [1] only considers a single object without clutter, each measurement is assumed to originate from the target. Starting with these assumptions, a tracking algorithm for an extended object or group of objects is derived as a recursive updating scheme for the conditional probability density

$$p(x_t, X_t|Z^t) = p(x_t|X_t, Z^t) p(X_t|Z^t) = \mathcal{N}(x_t|\bar{x}_{t|t}, P_{t|t} \otimes X_t) ZW(X_t; \zeta_{t|t}, X_{t|t}) \quad (1)$$

at each scan $t$, where $\otimes$ designates the Kronecker product as it is introduced in Appendix A. $P_{t|t} \in \mathbb{R}^{s \times s}$ is the estimation error covariance. $ZW$ is an inverted Wishart density and $\zeta_{t|t}$ is its scalar Wishart parameter (Appendix B). $\mathcal{N}$ denotes the Gaussian as usual. Decoupling the joint density into a product of a Gaussian for the kinematical state and an inverted Wishart density for the extent makes it possible to handle the two corresponding estimation problems separately. The choice of an inverted Wishart density to represent the uncertainty within the object extent, relies on the concept of conjugate priors in Bayesian inference. The fact that the product of a Wishart density (extent likelihood) and an inverted Wishart density (conjugate prior) results in an inverted Wishart density again, is the essential precondition for the derivation of a recursive filtering scheme.

A. Object and Measurement Model

The extended object is assumed to move according to the discrete-linear model

$$x_t = \Phi_{t|t-1} x_{t-1} + v_t \quad \text{with} \quad v_t \sim \mathcal{N}(0, \Delta_{t|t-1}) \quad , \quad (2)$$

where the evolution matrix is defined as $\Phi_{t|t-1} = F_{t|t-1} \otimes I_d$ and $\Delta_{t|t-1} = Q_{t|t-1} \otimes X_t$ is the process noise covariance that is controlled by the actual parameter matrix $Q$ and the object extent $X_t$. The $s \times s$ matrices $F_{t|t-1}$ and $Q_{t|t-1}$ in this form can also be found in the one-dimensional tracking problem.

In the measurement model, each of the $N_t$ measurements caused by the extended object is interpreted as an independent position measurement of the object centroid $x_t$ according to

$$z_t^i = Hx_t + w_t^i \quad \text{with} \quad w_t^i \sim \mathcal{N}(0, X_t) \quad . \quad (3)$$

This model is based on the assumption that the uncertainty of a single measurement is dominated by the object extent. The precision error within the measuring process (due to the sensor device) only reflects the position accuracy with respect to a certain scattering center on the object. But the larger the extent, the less important it is which of the scattering centers were actually responsible for the received measurements. The observation matrix is defined as $H = [h_1 1_d, h_2 1_d] = H \otimes I_d$, with $h_1 = 1$ and $h_2 = 0$ and $H = [1,0]$.

B. Prediction Equations

Under the described assumptions the following prediction equations can be derived for the kinematical state of the object centroid:

$$x_{t|t-1} = (F_{t|t-1} \otimes I_d)x_{t-1|t-1} \quad (4)$$

$$P_{t|t-1} = F_{t|t-1}P_{t-1|t-1}F_{t|t-1}^T + Q_{t|t-1} \quad (5)$$

Note that the covariance matrix $P_{t|t-1}$ of the estimation error is merely of dimension $s \times s$ and that the actual uncertainty $P_{t|t-1} \otimes X_t$ is also influenced by the object extent.

The prediction equations for the object extent are based on a heuristic temporal decay model with time-constant $\lambda$:

$$\zeta_{t|t-1} = \exp\left(\frac{-\delta_{t|t-1}}{\lambda}\right) \zeta_{t-1|t-1} \quad (6)$$

$$X_{t|t-1} = \frac{\zeta_{t-1|t-1} - \delta_{t|t-1}}{\zeta_{t-1|t-1} - \delta_{t|t-1}} X_{t-1|t-1} \quad , \quad (7)$$

so that the exponential term in Eq. (6) controls to what degree the already assimilated information from past scans has an
impact on the estimate of the present scan. In that sense the
exponential term serves as a flexibility factor, where $\lambda = \infty$
corresponds to a static object, which means that all information
from the past is preserved.

C. Filtering Equations

In the filtering step the currently received $N_t$ measurements
are processed. The likelihood function of the Bayesian single
extended target filter is given by

$$p(Z_t|N_t, x_t, X_t) \propto \mathcal{N}(\bar{z}_t; Hx_t, \frac{X_t}{N_t}) \mathcal{W}(\hat{Z}_t; N_t - 1, X_t),$$
(8)

with a measurement centroid and a spread matrix defined as

$$\bar{z}_t = \frac{1}{N_t}\sum_{r=1}^{N_t} z_r^i$$
and
$$\hat{Z}_t = \sum_{r=1}^{N_t}(z_r^i - \bar{z}_t)(z_r^i - \bar{z}_t)^T.$$  
(9)

With this likelihood function the following filtering equations
can be developed to update the cinematic state of the target
centroid and the corresponding covariance matrix:

$$x_{t|t} = x_{t|t-1} + (W_{t|t-1} \otimes I_d)(\bar{z}_t - Hx_{t|t-1})$$
(10)

$$P_{t|t} = P_{t|t-1} - W_{t|t-1}S_{t|t-1}W_{t|t-1}^T,$$  
(11)

with the gain matrix $W_{t|t-1}$ and the innovation factor $S_{t|t-1}$
defined in close analogy to the standard Kalman filtering:

$$W_{t|t-1} = P_{t|t-1}^{-1}H^TS_{t|t-1}^{-1}$$
(12)

$$S_{t|t-1} = \bar{H}P_{t|t-1}\bar{H}^T + \frac{1}{N_t}.$$  
(13)

The target extent matrix and the Wishart parameter are estimated
by summation over time:

$$X_{t|t} = X_{t|t-1} + Y_{t|t-1} + \hat{Z}_t,$$
(14)

$$\zeta_{t|t} = \zeta_{t|t-1} + N_t,$$  
(15)

with an innovation matrix $Y_{t|t-1}$ given by

$$Y_{t|t-1} = S_{t|t-1}^{-1}(\bar{z}_t - Hx_{t|t-1})(\bar{z}_t - Hx_{t|t-1})^T.$$  
(16)

Note that we are interested in the expectation of the extent
and that this is the fraction of the random matrix estimate and
the parameter estimate according to Appendix B. Retroduction
equations are presented in [1] but are not applied in this work.

IV. STANDARD HISTOGRAM PMHT

The histogram PMHT, as introduced in [2], is now reviewed.
The H-PMHT is derived by interpreting the sensor image as
a histogram of observations of an underlying random process.
The received energy in each cell is quantized, and the resulting
integer is treated as a count of the number of measurements
that fell within that cell. The sum over all of the cells is the
total number of measurements taken. The probability mass
function for these discrete measurements is modeled as a
multinomial distribution where the probability mass for each
cell is the superposition of target and noise contributions.

The probability that an individual histogram shot falls in
cell $\ell$ is

$$p_{t}(X_{\ell}) = \frac{p_{t}(X_{\ell})}{p(X_{t})}, \quad \text{where} \quad p(X_{t}) = \sum_{\ell} p_{t}(X_{\ell})$$

and $p_{t}(X_{\ell}) = \int_{B_{\ell}} f(\tau|X_{\ell})\,d\tau$.  
(17)

The spatial extent of cell $\ell$ (of arbitrary dimensionality) is $B_{\ell}$
and $X_{\ell} = [x_{\ell}^1, \ldots, x_{\ell}^M]$ is the system state vector at time $\tau$, 
i.e. it summarizes all of the targets. The underlying continuous
spatial density, $f(\tau|X_{t})$ is the superposition of a background
clutter model ($m = 0$) and $M$ target components

$$f(\tau|X_{\ell}) = \sum_{m=1}^{M} \pi_{\ell}^{m} G^{m}(\tau|x_{\ell}^{m}),$$  
(18)

where the vector $x_{\ell}^{m}$ is the state of the target $m$ at time $\tau$, and
the mixing proportions form a probability vector, i.e. $\pi_{\ell}^{m} \geq 0$
and $\sum_{m=0}^{M} \pi_{\ell}^{m} = 1$. The mixing proportions $\Pi_{t}$ can be
interpreted as the relative power of each target. In the simplest
case, the background clutter model is uniform and $G^{0}(\tau)$ is a constant.
In spatially non-uniform clutter, mapping approaches may be used.
The component $G^{m}>0(\tau|x_{\ell}^{m})$ may be common
across targets, e.g. it could represent the point spread function
for a sensor observing point scatterers, or each target may
have a unique signature, such as with high resolution optical
sensors.

The H-PMHT is an EM algorithm that treats the assignment
of histogram shots to the model components and the precise
location of shots as missing data. In addition, it allows for
unobserved cells that are notionally sensor pixels for which
no data was collected. The data from these unobserved cells is
also treated as missing data. Assuming an existing estimate of
the state $X_{t}^{(i)}$ and the mixing proportions $\Pi_{t}^{(i)}$ at iteration $i$, the
H-PMHT algorithm determines the probability of the missing
data in the expectation step (E-step) and then refines the state
estimates in the maximization step (M-step). The probability
of the missing data defines an auxiliary function $Q$ that can be
interpreted as an optimal lower bound of the objective function
$p_{t}(X_{t}|Z_{1}^{(i)}, \ldots, Z_{T})$ that actually is to be maximized
to solve the multi-target tracking problem. The bound is optimal
in the sense that it touches the objective function at the current
estimate. In the M-Step, this bound is maximized with respect
to the parameters to be estimated.

A. Expectation Step

We now introduce the equations to calculate the probability
of the missing data. The per-cell proportion of the contribution
of target $m$ at scan $t$ (target cell probability) is denoted as

$$P_{t}^{m(i)} := p_{t}(X_{t}^{m(i)}) = \int_{B_{\ell}} G^{m}(\tau|x_{t}^{m(i)})\,d\tau.$$  
(19)

Then the total cell probabilities are given by

$$P_{t}^{(i)} := p_{t}(X_{t}^{(i)}) = \sum_{m=0}^{M} \pi_{\ell}^{m(i)} p_{t}^{m}(X_{t}^{m(i)}).$$  
(20)
The observed power in cell $\ell$ at scan $t$ is denoted $z^\ell_t$. Then the expected measurement $\bar{z}^\ell_t$ in cell $\ell$ is given by
\[
\bar{z}^\ell_t = \left\{ \begin{array}{ll} z^\ell_t & \ell \in L, \\
\|Z\| \frac{P^\ell_t(p^{G(\tau)}_t)}{P^\ell_t(p^{f(i)}_t)} & \ell \notin L \quad \text{with} \quad \|Z\| = \sum_{\ell \in L} z^\ell_t,
\end{array} \right.
\]
(21)
where $L$ is the set of all observed cells and $\bar{L}$ is the set of all unobserved cells that may be empty. In the following, $S$ denotes the union $L \cup \bar{L}$ of the cell sets. For abbreviation we use the calligraphic notation $P$ without the state variable. $P^{G(\tau)}_t$ denotes the total sensor probability $P^{f(i)}_t = \sum_{\ell \in \ell} P^{f(i)}_t$.

**B. Maximization Step**

The parameters to be maximized during the M-step are the mixing proportions $\pi_t^m$ and the kinematical states $x_t^m$ for all targets at all scans. The improved mixing proportion estimate is given by
\[
\pi_t^{m(i+1)} = \frac{\pi_t^{m(i)} V_t^m}{\sum_{m=0}^M \pi_t^{m(i)} V_t^m},
\]
(22)
with the placeholder
\[
V_t^m(i) := \sum_{\ell \in S} \bar{z}^\ell_t \left( \frac{P_t^f(p^{f(i)}_t)}{P_t^f(p^{G(i)}_t)} \right),
\]
(23)
that will be frequently used in the discussion that follows.

The state sequence for each target is estimated independently by maximizing the function
\[
Q^m = \sum_{t=1}^T \|Z\| \log \left\{ \frac{p(x_t^m|x_{t-1}^m)}{p_t^f(p^{f(i)}_t)} \right\} +
\sum_{t=1}^T \sum_{\ell \in S} \pi_t^{m(i)} \bar{z}^\ell_t \int_{B_t} G^m(\tau | x_t^m) \log G^m(\tau | x_t^m) d\tau
\]
(24)
For the case of linear Gaussian statistics, Streit demonstrates in [2] that this maximization problem is equivalent to a point measurement filtering problem with synthetic measurements
\[
\bar{z}^{f(i)}_t = \frac{1}{\nu_t^m(i)} \int_{B_t} G^m(\tau | x_t^m) d\tau
\]
(25)
where the cell-level centroid $\bar{z}^{f(i)}_t$ is given by
\[
\bar{z}^{f(i)}_t = \frac{1}{\nu_t^m(i)} \int_{B_t} G^m(\tau | x_t^m) d\tau
\]
(26)
The associated synthetic measurement covariance is
\[
\bar{R}^{f(i)}_t = \frac{R}{\pi_t^{m(i)} V_t^m}
\]
(27)
and the synthetic process covariance is
\[
\bar{Q}^{f(i)}_t = \frac{1}{\|Z\|} \bar{Q}^{f(i)}_t
\]
(28)
Details about the derivation of H-PMHT can be found in [2]. A Kalman smoother can be used to solve this point measurement filtering problem and to calculate improved target states.

**C. Standard H-PMHT Algorithm**

After proper initiation, the H-PMHT algorithm consists of iteratively repeating the following steps for each scan $t$ of a batch of data from $t = 1$ up to $t = T$:

1) compute cell probabilities, Eq. (19), (20)
2) ——— expected measurements, Eq. (21)
3) ——— cell-level centroids, Eq. (26)
4) ——— synthetic measurements, Eq. (25)
5) ——— synthetic covariance matrices, Eq. (27), (28)
6) ——— mixing proportion estimates, Eq. (22)
7) apply a Kalman smoother with synthetic measurements to get improved target state estimates
8) if $Q$-function has converged shift the batch and start a new iteration process, otherwise go back to 1)

In this work we apply a batch of length $T = 1$ and therefore replace the smoother by a Kalman filter. The convergence is checked by considering the change of the target state estimates, since a calculation of the $Q$-function takes too much time.

In the original derivation of H-PMHT, Streit also proposed an estimator for the covariance matrix $R$ [2], which actually represents the target extent. In this case the fixed matrix $R$ is replaced by $\bar{R}^{f(i)}_t$, a sum of weighted spread matrices:
\[
\bar{R}^{f(i)}_t = \frac{1}{\nu_t^m(i)} \sum_{\ell \in S} \bar{z}^\ell_t \left( \frac{P_t^f(p^{f(i)}_t)}{P_t^f(p^{G(i)}_t)} \right) \bar{R}^{f(i)}_t
\]
(29)
with
\[
\bar{R}^{f(i)}_t = \int_{B_t} G^m(\tau | x_t^m) \left( \tau - Hx_t^m \right) \left( \tau - Hx_t^m \right)^T d\tau
\]
(30)
Note that these extent estimates are independent from scan to scan and are not governed by a recursive process as in [1].

**V. Histogram PMHT with Random Matrices**

In this section we incorporate random matrices as additional state variables into the framework of histogram Probabilistic Multi-Hypothesis Tracking (H-PMHT). The derived algorithm unifies the concepts in Section III and IV and represents a novel approach to track multiple extended objects using image data. Following [1], we assume that the target extents can be modeled as ellipsoids, which allows their representation by random matrices that are to be estimated simultaneously with the kinematical states. Data assignment conflicts with clutter or other targets are resolved within the PMHT. Moreover, the histogram-based version of PMHT incorporates the ability of track-before-detect (TkBD) into the algorithm and makes it able to extract extended object tracks directly from an image sequence. We consider the linear Gaussian case.

**A. Statement of the Auxiliary Function**

The statement of the EM auxiliary function ($Q$-function) is essential in the derivation of any EM-based approach. The integration of recursively estimated target extents into the H-PMHT framework requires the introduction of the random variable $X_t^m$ as an additional variable for each target at each
scan. The derivation is carried out in 6 steps. Eq. (33) shows the Q-function with additional random matrices, which is the starting point (step 1). This expression is analogous to Eq. (24). We assume that all densities $G^m$ are Gaussians and write $G$.

**B. Expectation Step**

To determine the Q-function, the probability of the missing data given the current estimates has to be calculated in the E-step. The target cell probabilities are now given by

$$P_t^{(m(i))} := p_t^m \left( x_t^{m(i)} \right) = \int_{B_t} N \left( n \tau, H x_t^{m(i)}, X_t^{m(i)} \right) d \tau \,, \quad (31)$$

so that the weight of each cell is governed by $X_t^{m(i)}$, i.e. the expectation of the extent (Appendix B). In case of a relevant measurement error, $X_t^{m(i)}$ could be replaced by $X_t^{m(i)} + R$. However, it is important to emphasize that this is a heuristic, since the measurement error is not taken into account by the Bayesian extended object filter that is exploited in the M-step (see Section III-A). The definition of the total cell probability $P_t^{(i)}$ and the total sensor probability $P_t^{(i)}$ is straightforward.

**C. Maximization Step**

In the M-step the auxiliary function $Q$ has to be maximized with respect to the free variables $x_t^m, X_t^m$ and $p_t^m$, for each target at each scan. The mixing proportions $\pi_t^m$ are obtained according to Eq. (22). To maximize the extended target states some more work is required.

Starting from Eq. (33) we first focus on the likelihood function $L(Z, X^m)$, where $X^m$ refers to the extended target states. The logarithm of $G$ leads to two terms: one from the normalizing term and one from the exponent, giving Eq. (34). The exponent term, denoted $\Lambda(Z_t, X^m)$, is expanded and the linear expectations are replaced by $\bar{Z}_t^{m(i)} = \mu_t^{m(i)}$, the synthetic measurement definition in Eq. (25), leading to Eq. (35). At this stage it is already possible to divide the terms into those that belong to the kinematic likelihood and those that are part of the extent. But there are some terms missing and have to be added to complete the quadratic terms. We choose the extension

$$\Lambda(Z_t, X^m) = \ldots \pm 2 \pi_t^m \mu_t^{m(i)} \left( \bar{Z}_t^{m(i)} - \mu_t^{m(i)} \right)^T X_t^m \left( \bar{Z}_t^{m(i)} - \mu_t^{m(i)} \right)$$

leading to 4 additional terms. Then, one of the positive terms becomes part of the kinematic likelihood that now looks as usual in H-PMHT, except that the measurement covariance is replaced by the extent matrix (see Eq. (36)). The remaining three terms comprise the expression $\bar{Z}_t^{m(i)} = \mu_t^{m(i)}$ that is now re-substituted according to the synthetic measurement definition in Eq. (25). Thus, there are 4 terms each with an integral over a Gaussian that can be combined to a quadratic term as shown in the second row of Eq. (36) representing the extent part. Substituting $\Lambda(Z_t, X^m)$ in the likelihood function, and applying the algebraic rule in Eq. (43) followed by the exponential yields Eq. (37). This equation contains the definition of a spread matrix $\bar{Z}_t^{m(i)}$ that is built analogously to the usual synthetic measurements in Eq. (25). From there, it can be easily seen that the extent part can be represented by a Wishart density (Appendix B), whereas the kinematic part is a Gaussian (see equations in Appendix A). This leads to the final likelihood function of H-PMHT with extent estimates using random matrices in Eq. (38). Note the similarity between the single extended target likelihood in Eq. (8) and the multi-target likelihood in Eq. (38). Obviously the number of measurements $N_i$ in the single target version is now replaced by $\pi_t^m \nu_t^m$ denoting the relative target power times the weighted sum of cell powers. This is exactly the amount of measurement information that is assigned to the respective target after the data assignment phase.

Finally, the evolution part $E(X'^m)$ of Eq. (33) has to be developed. In [1] it has been shown that under certain assumptions the predicted density $p(x_t^m, X_t^m|Z^{t-1})$ can be obtained by two independent integrations. We follow these assumptions and apply the logarithm in the evolution part. Then, with the model in Section III-A the updated process noise covariance matrix

$$\Delta_t^{m(i+1)} = \frac{P_t^{(i)} Q_t^{(i-1)}}{\| Z_t \|} \otimes X_t^{m(i+1)}$$

can be derived based on the same method as in standard H-PMHT. The state sequence of an extended target $m$ is thus optimized by using the Bayesian extended target filter in [1] with synthetic data.

**Remark:** Let us consider the novel H-PMHT with only a single target component ($M = 1$, no clutter). In this case

$$\pi_t^{1(i)} \nu_t^{1(i)} = 1 \times \sum_{i \in S} \bar{Z}_t^{i} \equiv N_t$$

which is independent of $i$. If we set $\bar{Z}_t^{i}$ to 1 in case the power of the cell exceeds a certain threshold, and 0 otherwise, then the scenario represents a point measurement scenario with $N_t$ measurements, each of them denoting a position in the field of view. Thus, the Bayesian extended target filter in [1] is a special case of the proposed H-PMHT with random matrices.

**D. H-PMHT Algorithm with Random Matrices**

The data flow of the algorithm corresponds to the standard H-PMHT as introduced in Section IV-C with the following replacements: target cell probabilities are calculated according to Eq. (31), the synthetic measurement covariance matrix is replaced by a spread matrix according to Eq. (37), the process noise covariance is estimated according to Eq. (39) and the maximization is done by applying the Bayesian extended target filter in [1] for each target with synthetic data.

Since the PMHT itself is not able to extract and delete tracks, the novel H-PMHT also requires to be embedded into a track management system (TMS). This TMS holds two types of tracks: established tracks that have been accepted as target tracks and tentative tracks that are not confirmed. First, the established tracks are updated and tested to find which should be terminated. Then, the residual sensor image, i.e. all cell data that has not been assigned to the established tracks, is used to update the tentative tracks. After erasing the cell data used for the
Step 1: Q-function term referring to the state sequence of a single extended object track \([x, x_t]^{T}_{t=1}
\]

\[
Q_X = \sum_{t=1}^{T} \frac{\|Z_t\|}{P_t^{(i)}} \log \left\{ p\left( x_t^m | x_t^{m-1}, x_t \right) p\left( x_t^{m-1} \right) \right\} + \sum_{t=1}^{T} \frac{\pi_t^{m(i)} z_t^{m(i)}}{P_t^{(i)}} \int_{B_t} G\left( \tau | x_t^{m(i)}, x_t^m \right) \log G\left( \tau | x_t^{m(i)}, x_t^m \right) d\tau
\]

Step 2: Application of the logarithm followed by the exponential of the Gaussian

\[
\mathcal{L}(Z, X^m) = \sum_{t=1}^{T} \frac{\pi_t^{m(i)} z_t^{m(i)}}{P_t^{(i)}} \int_{B_t} \mathcal{N}\left( \tau; Hx_t^{m(i)}, x_t^{m(i)} \right) \left( \tau - Hx_t^{m(i)} \right)^T [X_t^m]^{-1} \left( \tau - Hx_t^{m(i)} \right) d\tau
\]

Step 3: Substitution of synthetic measurement expressions

\[
\Lambda(Z_t, X^m) = \sum_{t=S}^{T} \frac{\pi_t^{m(i)} z_t^{m(i)}}{P_t^{(i)}} \int_{B_t} \mathcal{N}\left( \tau; Hx_t^{m(i)}, x_t^{m(i)} \right) \tau^T [X_t^m]^{-1} \tau d\tau
\]

Step 4: Completion of the square and re-substitution of synthetic measurement expressions

\[
\Lambda(Z_t, X^m) = \left( \overline{z}_t^{m(i)} - Hx_t^m \right)^T \left[ \frac{X_t^m}{\pi_t^{m(i)} z_t^{m(i)}} \right]^{-1} \left( \overline{z}_t^{m(i)} - Hx_t^m \right)
\]

Step 5: Substitution of the \(\Lambda(Z_t, X^m)\) terms, application of the algebraic rule in Eq. (43) and the exponential

\[
\mathcal{L}(Z, X^m) \propto \prod_{t=1}^{T} \left[ 2\pi X_t^m \right]^{-1} \frac{1}{2} \pi_t^{m(i)} z_t^{m(i)} \times \text{etr} \left( -\frac{1}{2} \left( \overline{z}_t^{m(i)} - Hx_t^m \right)^T \left[ \frac{X_t^m}{\pi_t^{m(i)} z_t^{m(i)}} \right]^{-1} \left( \overline{z}_t^{m(i)} - Hx_t^m \right) \right)
\]

Step 6: Final likelihood function of H-PMHT with target extent estimates

\[
\mathcal{L}(Z, X^m) \propto \mathcal{N}\left( \overline{z}_t^{m(i)}; Hx_t^m, \left[ \frac{X_t^m}{\pi_t^{m(i)} z_t^{m(i)}} \right] \right) \mathcal{W}\left( \left( \overline{z}_t^{m(i)}; \pi_t^{m(i)} z_t^{m(i)} \right), -1, X_t^m \right)
\]
tentative tracks, the final residual image is used to initiate new
tentative tracks. This is done by applying a peak detection
algorithm and then combining the peaks of two successive
scans via two-point differencing. Note that the peak detector
is solely applied to get the seeds for new tentative tracks.
The whole track maintenance of both established and tentative
tracks is completely performed by the H-PMHT, without any
thresholding. To check the validity of each track, the estimated
mixing proportions are used to derive an SNR according to
\[ \text{SNR}_i^n = 20 \log_{10} \left( \frac{\pi^{\text{tr}}_i \| Z_i \|}{3\pi^{\text{tr}}_i} \right) \]  

VI. EXPERIMENTAL RESULTS

The H-PMHT with random matrices is now evaluated by
means of a simulated image sequence. Each frame consists of
an image with 100 × 100 cells showing a single moving target
in Rayleigh noise of standard deviation 1. The target extent
is modeled as a Gaussian shape of width 2 cells and length
10 cells, and with a maximum value of 1. To get the received
energy at each point of the surface, the shape is scaled by the
target amplitude. The amplitude is set to 2.5, corresponding
to 4 dB. Each frame is a superposition
of noise and target data. The target velocity is 1 cell/sec.
The measurement interval is \( \delta_{\text{hit}} = 1 \) sec. Fig. 1 shows the
first simulated data frame of an example sequence on the
left and the true target movement on the right. The duration
of each sequence is 70 scans. A process noise \( \nu_i = 1.0 \) is
applied to the true positions (red line). The disturbed positions
are shown as blue crosses. The target ellipse is exemplarily
drawn at three scans. The situation is challenging, since the
target is making a turn and the signal amplitude is weak.
We generated 500 simulations. Each sequence was processed
by the following trackers: H-PMHT with random matrices
(H-RM), H-PMHT with independent extent estimates using
Eq. (29) (H-IE), standard H-PMHT for point targets (H-PT),
standard (point measurement) PMHT and PDAF.

The PMHT approaches were all applied with a batch length
of 1 scan. For H-RM and H-IE the initial target extent was set
to \( X_0^{\text{tr}} = 2.0^2 I_2 \) with \( \zeta_0^{\text{tr}} = 4 \) for each new tentative track. All
other approaches were run with a fixed extent matrix \( R \). For the
point measurement approaches PMHT and PDAF, a peak
detection algorithm was applied at each scan to extract the
points to be processed. We used the discretized continuous-
time kinematic model.

Table I shows the tracking results for all approaches. The
the histogram PMHT with random matrices (H-RM) that does not require any further adjustments. It generates approximately

![Figure 1. First frame of simulated image data (left) and true target movement (right). The red line represents the trajectory. The blue crosses show the target positions after application of process noise. Extent ellipses are exemplarily drawn at three scans. The target amplitude corresponds to 4 dB. The target starts at cell [20, 55]. The axis lengths of the extent ellipse are [2, 10].](image)

![Figure 2. Position RMSE over all non-lost estimates. The approaches are distinguished by a combination of color and line style.](image)

![Figure 3. Estimated length of major semi-axis and angle. The blue solid line shows the truth. The black line corresponds to the results of H-RM, while the dashed red line represents the H-IE.](image)

<table>
<thead>
<tr>
<th>Alg.</th>
<th>Total Lost</th>
<th>True</th>
<th>Lost</th>
<th>True</th>
<th>Adjustments</th>
</tr>
</thead>
<tbody>
<tr>
<td>H-RM</td>
<td>1.014</td>
<td>0.022</td>
<td>0.042</td>
<td>0.897</td>
<td>none</td>
</tr>
<tr>
<td>H-IE</td>
<td>1.014</td>
<td>0.015</td>
<td>0.027</td>
<td>0.921</td>
<td>none decrease ( R )</td>
</tr>
<tr>
<td>H-PT</td>
<td>1.014</td>
<td>0.015</td>
<td>0.027</td>
<td>0.921</td>
<td>( R = 2.0^2 I_2 ) decrease ( \pi_0 )</td>
</tr>
<tr>
<td>PMHT</td>
<td>1.040</td>
<td>0.036</td>
<td>0.066</td>
<td>0.924</td>
<td>( R = 2.0^2 I_2 ) replace ( \pi_0 )</td>
</tr>
<tr>
<td>PDAF</td>
<td>1.006</td>
<td>0.048</td>
<td>0.063</td>
<td>0.925</td>
<td>( R = 5.5^2 I_2 )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Alg.</th>
<th>Total Lost</th>
<th>True</th>
<th>Lost</th>
<th>True</th>
<th>Adjustments</th>
</tr>
</thead>
<tbody>
<tr>
<td>H-RM</td>
<td>1.014</td>
<td>0.022</td>
<td>0.042</td>
<td>0.897</td>
<td>none</td>
</tr>
<tr>
<td>H-IE</td>
<td>1.014</td>
<td>0.015</td>
<td>0.027</td>
<td>0.921</td>
<td>none decrease ( R )</td>
</tr>
<tr>
<td>H-PT</td>
<td>1.014</td>
<td>0.015</td>
<td>0.027</td>
<td>0.921</td>
<td>( R = 2.0^2 I_2 ) decrease ( \pi_0 )</td>
</tr>
<tr>
<td>PMHT</td>
<td>1.040</td>
<td>0.036</td>
<td>0.066</td>
<td>0.924</td>
<td>( R = 2.0^2 I_2 ) replace ( \pi_0 )</td>
</tr>
<tr>
<td>PDAF</td>
<td>1.006</td>
<td>0.048</td>
<td>0.063</td>
<td>0.925</td>
<td>( R = 5.5^2 I_2 )</td>
</tr>
</tbody>
</table>

\( ^1 \text{LT} = \otimes \text{length of true tracks}, \ ^2 \text{BT} = \otimes \text{birth time of all tracks} \)
one track per run and has the best track loss statistics. All other approaches were significantly worse. For PDAF and PMHT the chosen measurement covariance matrix \( \mathbf{R} \) was too small. The H-IE did not trust enough into the target data, since the extent approaches were significantly worse. For PDAF and PMHT the approach achieves better tracking results than the PDAF, the standard PMHT, the standard H-PMHT and the H-PMHT with independent extent estimates. Moreover, the Bayesian extended object filter in [1] based on random matrices was shown to be a special case of H-PMHT with random matrices.

### Appendix A

**Some Facts from Linear Algebra**

The Kronecker product \( \otimes \) of two matrices \( \mathbf{A} = (a_{ij})_{i=1,j=1}^{m,n} \) and \( \mathbf{B} = (b_{ij})_{i=1,j=1}^{p,q} \) is defined as

\[
\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix}
    a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\
    a_{21}\mathbf{B} & a_{22}\mathbf{B} & \cdots & a_{2n}\mathbf{B} \\
    \vdots & \vdots & \ddots & \vdots \\
    a_{m1}\mathbf{B} & a_{m2}\mathbf{B} & \cdots & a_{mn}\mathbf{B}
\end{bmatrix}
\]

Furthermore, for column vectors \( \mathbf{x} \) and \( \mathbf{y} \) of equal dimension the following equations hold:

\[
\mathbf{x}^T \mathbf{y} = \text{tr}(\mathbf{x} \mathbf{y}^T), \quad \mathbf{x}^T \mathbf{A}^{-1} \mathbf{x} = \text{tr}(\mathbf{x} \mathbf{x}^T \mathbf{A}^{-1}) = \text{tr}(\mathbf{A}^{-1} \mathbf{x} \mathbf{x}^T), \quad \mathbf{I} + \mathbf{x} \mathbf{y}^T = 1 + \mathbf{x}^T \mathbf{y}
\]

and

\[
\exp(-\frac{1}{2} \mathbf{x}^T \mathbf{A}^{-1} \mathbf{x}) = \text{etr}(-\frac{1}{2} \mathbf{x} \mathbf{x}^T \mathbf{A}^{-1}),
\]

with \( \text{etr}(\mathbf{A}) \) representing \( \exp(\text{tr}(\mathbf{A})) \). For proofs see e.g. [8].

### Appendix B

**Wishart Densities**

A \( d \times d \) SPD random matrix \( \mathbf{X} \) is Wishart-distributed, if its density is given by [7, p. 87] \((m \geq d)\):

\[
W(\mathbf{X}; m, \mathbf{C}) \propto |\mathbf{X}|^{-\frac{m+d-1}{2}} \text{etr}(-\frac{1}{2} \mathbf{C}^{-1} \mathbf{X})
\]

with a scalar parameter \( m \) and a \( d \times d \) SPD matrix \( \mathbf{C} \). Its expectation is given by \( E[\mathbf{X}] = m \mathbf{C} \), with \( \mathbf{C} = (c_{ij}) \) and \( \mathbf{X} = (x_{ij}) \) [7, p. 98]. A \( d \times d \) SPD random matrix \( \mathbf{X} \) is inverted-Wishart-distributed, if its density is given by [7, p. 111] \((m > 2d)\):

\[
D W(\mathbf{X}; m, \mathbf{C}) \propto |\mathbf{X}|^{-\frac{m+d-1}{2}} \text{etr}(-\frac{1}{2} \mathbf{X}^{-1} \mathbf{C}^{-1})
\]

with a scalar parameter \( m \) and a \( d \times d \) SPD matrix \( \mathbf{C} \). Its expectation is given by \( E[\mathbf{X}] = \frac{m \mathbf{C}}{m-d} \) for \((m-d > 1)\), with \( \mathbf{C} = (c_{ij}) \) and \( \mathbf{X} = (x_{ij}) \) [7, p. 113].

### References


