

Out-of-Sequence processing of cluttered sensor data using multiple evolution models

Felix Govaers

Sensor Data and Information Fusion
and Institute of Computer Science 4
Fraunhofer FKIE and University of Bonn
Wachtberg and Bonn, Germany
govaers@cs.uni-bonn.de

Wolfgang Koch

Sensor Data and Information Fusion
Fraunhofer FKIE
Wachtberg
wolfgang.koch@fkie.fraunhofer.de

Abstract – *In target tracking applications, the full information on the kinematic target states accumulated over a certain time window up to the present time is contained in the joint probability density function of these state vectors, given the time series of all sensor data. In [1] the structure of this Accumulated State Density (ASD) has been revealed. Furthermore, ASDs enable us to process Out-of-Sequence (OoS) measurements in a neat and straightforward way. This paper presents an algorithm for the processing of OoS measurements in situations with more relaxed assumptions. On the one hand, sensors often return ambiguous measurement data. Then, measurement association methodologies as the Multi-Hypothesis Tracker (MHT) are required. On the other hand, the evolution model in use might not be unique. The well-known approach to this challenge is the Interacting Multiple Model (IMM) filter. In this paper, an IMM/MHT extension to the ASD paradigm is discussed, tested by simulation, and evaluated.*

1 Introduction

1.1 Motivation

The problem of Out-of-Sequence (OoS) measurements is inherent in many applications of distributed sensors. *Best-effort* protocols for a network transport layer are able to deliver data almost in real-time, but often effects like re-orderings or even burst re-orderings of packets can be observed. These phenomena appear e.g. by multi-path packet delivery in networks or by bottle-neck links in an inhomogeneous network topology.

In [1] it becomes clear that the information necessary for a neat processing of OoS measurements is contained in the joint density of all states within a certain time window, which is at least as big as the time lag of the arriving measurement. This is similar to the results of [13] where a pseudo innovation of the corresponding lag is calculated and then fused with the delayed data. This lag may not be greater than a given parameter, which is part of the modeling assumption. However, these considerations can be generalized to the case of ambiguity with respect to the origin of the sensor data and with respect to the evolution model, cur-

rently in effect, i.e. to Multiple Hypothesis Tracking (MHT) and Interacting Multiple Model filters (IMM).

The paradigm presented in this paper aims at processing OoS measurements on the fly even if clutter is present and the system model in use might not be unique. To the authors knowledge, this is the first time an algorithm is presented and evaluated which combines a Multi-Hypothesis/Interacting Multiple Model (MHT/IMM) tracker with a filter methodology to process OoS measurements. To this end, we constitute so called *Accumulated State Densities (ASD)*, which were introduced by W. Koch in [1]. These contain the correlation information of all states within a given time window and enable us to calculate the impact of a measurement on all these states.

Structure This paper is organized as follows. The current section ends with a clear and more formal formulation of the problem. Section 2 summarizes the structure and the Bayesian predictor-corrector methodology for filtering of the ASDs. Then, in section 3 this framework is extended to the paradigms of MHT/IMM. In the last section, we present a discussion and an evaluation of simulation results.

1.2 Formulation of the problem

The knowledge of state quantities as a vector \mathbf{x}_l at arbitrary time t_l given all sensor data Z^k up to time t_k can be represented by the conditional *probability density function (pdf)* $p(\mathbf{x}_l|Z^k)$. This is why many state estimation algorithms calculate or approximate this pdf [3, 4]. In order to allow OoS measurements as input to a Bayesian estimator in a natural way, it is crucial to investigate the properties and correlations between each state. This can be done by calculating the joint pdf $p(\mathbf{x}_k, \dots, \mathbf{x}_n|Z^k)$ of the accumulated state vector $\mathbf{x}_{k:n} := (\mathbf{x}_k^\top, \dots, \mathbf{x}_n^\top)^\top$ on a certain time window $[t_n, t_k]$, which corresponds to $k-n$ steps. Whenever the standard Kalman filter is applicable (perfect sensor-data-to-track association, linear Gaussian sensor and evolution models), it is possible to calculate this *accumulated state density (ASD)* as is shown in [1] and sketched in the next section. As long as the time stamp of an arriving measurement is

within that time window, an update of the accumulated state can be calculated independent of whether it is in-Sequence or Out-of-Sequence.

In this paper, we show that the postulation of perfect sensor-data-to-track association can well be relaxed. Therefore, let us assume that the output of a sensor at time t_l is given as a set of measurements $\{\mathbf{z}_l^{j_i}\}_{j_i=1,\dots,q_l}$, where the number of measurements q_l may vary in time. Furthermore, if the sensor's detection probability P_D is smaller than 1, it might happen that all measurements $\mathbf{z}_k^{j_i}$ of one scan are clutter.

The other extension to the ASD filtering algorithm involves multiple evolution models. In practical applications, it may be uncertain which evolution model out of a set of r possible alternatives is currently in effect (different flight phases such as no turn, slight maneuver, high- g turn, e.g.). The maneuvering class $1 \leq i_k \leq r$ an object belongs to at time t_k can thus be considered as a part of its state. The presented work states the prediction and update equations under consideration of multiple models.

2 State estimation with Accumulated State Density Filters

This section summarizes the prediction and update step of an ASD filter. For shortening reasons we have to skip the full derivation of the ASD covariance matrix which can be read in detail in [1].

At first, assume a scenario where the Kalman filter is applicable. As an induction start, let the state density for all sensor data up to time t_{k-1} be given by a Gaussian pdf:

$$p(\mathbf{x}_{k-1:n}|Z^{k-1}) = \mathcal{N}(\mathbf{x}_{k-1:n}; \mathbf{x}_{k-1:n|k-1}, \mathbf{P}_{k-1:n|k-1}). \quad (1)$$

Here, $\mathbf{x}_{k-1:n} = (\mathbf{x}_{k-1}^\top, \dots, \mathbf{x}_n^\top)^\top$ is the random vector for all $k-n$ states at t_{k-1}, \dots, t_n respectively. Analogously, $\mathbf{x}_{k-1:n|k-1} = (\mathbf{x}_{k-1|k-1}^\top, \dots, \mathbf{x}_n^\top)^\top$ is the vector consisting of all expectation values of a Kalman filter with an application of the Rauch-Tung-Striebel (RTS) recursion [8], which yields the optimal single state pdfs $p(\mathbf{x}_l|Z^{k-1}) = \mathcal{N}(\mathbf{x}_l; \mathbf{x}_{l|k-1}, \mathbf{P}_{l|k-1})$ for $l \in \{k-1, \dots, n\}$. The accumulated state covariance $\mathbf{P}_{k-1:n|k-1}$ has the corresponding covariance matrices $\mathbf{P}_{k-1|k-1}, \dots, \mathbf{P}_{n|k-1}$ as its block diagonal entries. However, in general it is not block diagonal [1].

2.1 ASD prediction update

By renumbering the states it can be achieved that the given accumulated state $\mathbf{x}_{k-1:n}$ is extended to $\mathbf{x}_{k:n}$ such that $t_k \geq \dots \geq t_n$. Let the new measurement to be processed be \mathbf{z}_m with time stamp $t_m \in [t_n, t_k]$. I.e. \mathbf{z}_m is *In-Sequence*, if $t_m = t_k$ and *Out-of-Sequence*, if not.

By a repeated use of Bayes' rule and the Markov assumption, the ASD can be written as

$$p(\mathbf{x}_{k:n}|Z^{k-1}) \propto p(\mathbf{z}_k|\mathbf{x}_k) p(\mathbf{x}_k|\mathbf{x}_{k-1}) \cdots p(\mathbf{x}_m|\mathbf{x}_{m-1}) \cdots p(\mathbf{z}_{n+1}|\mathbf{x}_{n+1}) p(\mathbf{x}_{n+1}|\mathbf{x}_n) p(\mathbf{x}_n|Z^n) \quad (2)$$

As we assume a Gaussian linear evolution and sensor model given by

$$p(\mathbf{x}_l|\mathbf{x}_{l-1}) = \mathcal{N}(\mathbf{x}_l; \mathbf{F}_{l|l-1}\mathbf{x}_{l-1}, \mathbf{D}_{l|l-1}), \quad (3)$$

$$p(\mathbf{z}_l|\mathbf{x}_l) = \mathcal{N}(\mathbf{z}_l; \mathbf{H}_l\mathbf{x}_l, \mathbf{R}_l) \quad (4)$$

we may insert (3) and (4) now into (2):

$$p(\mathbf{x}_{k:n}|Z^{k-1}) \propto \prod_{l \neq m} [\mathcal{N}(\mathbf{z}_l; \mathbf{H}_l\mathbf{x}_l, \mathbf{R}_l) \cdot \mathcal{N}(\mathbf{x}_l; \mathbf{F}_{l|l-1}\mathbf{x}_{l-1}, \mathbf{D}_{l|l-1})] \cdot \mathcal{N}(\mathbf{x}_n; \mathbf{x}_{n|n}, \mathbf{P}_{n|n}) \cdot \mathcal{N}(\mathbf{x}_m; \mathbf{F}_{m|m-1}\mathbf{x}_{m-1}, \mathbf{D}_{m|m-1}), \quad (5)$$

where $\mathcal{N}(\mathbf{x}_n; \mathbf{x}_{n|n}, \mathbf{P}_{n|n})$ is the estimation density at time t_n for the state \mathbf{x}_n . Now, a repeated use of a well-known product formula (see (51) in the appendix) yields

$$p(\mathbf{x}_{k:n}|Z^{k-1}) \propto \mathcal{N}(\mathbf{x}_k; \mathbf{x}_{k|k}, \mathbf{P}_{k|k}) \prod_{l=n}^{k-1} \mathcal{N}(\mathbf{x}_l; \mathbf{h}_{l|l+1}\mathbf{x}_{l+1}, \mathbf{R}_{l|l+1}), \quad (6)$$

where for $l \neq m$ the functions $\mathbf{h}_{l|l+1}(\cdot)$ and covariances $\mathbf{R}_{l|l+1}$ are given by

$$\mathbf{h}_{l|l+1}(\mathbf{x}_{l+1}) = \mathbf{x}_{l|l} + \mathbf{W}_{l|l+1}(\mathbf{x}_{l+1} - \mathbf{x}_{l+1|l}), \quad (7)$$

$$\mathbf{R}_{l|l+1} = \mathbf{P}_{l|l} - \mathbf{W}_{l|l+1}\mathbf{P}_{l+1|l}\mathbf{W}_{m|m+1}^\top, \quad (8)$$

$$\mathbf{W}_{m|m+1} = \mathbf{P}_{m|m-1}\mathbf{F}_{m+1|m}^\top\mathbf{P}_{m+1|m}^{-1}, \quad (9)$$

and for $l = m$ the above equations become

$$\mathbf{h}_{m|m+1}(\mathbf{x}_{m+1}) = \mathbf{x}_{m|m-1} + \mathbf{W}_{m|m+1}(\mathbf{x}_{m+1} - \mathbf{x}_{m+1|m-1}), \quad (10)$$

$$\mathbf{W}_{m|m+1} = \mathbf{P}_{m|m-1}\mathbf{F}_{m+1|m}^\top\mathbf{P}_{m+1|m}^{-1}, \quad (11)$$

$$\mathbf{R}_{m|m+1} = \mathbf{P}_{m|m-1} - \mathbf{W}_{m|m+1}\mathbf{P}_{m+1|m}^{-1}\mathbf{W}_{m|m+1}^\top. \quad (12)$$

$$\mathbf{W}_{m|m+1}\mathbf{P}_{m+1|m-1}\mathbf{W}_{m|m+1}^\top. \quad (13)$$

This is due to the fact that \mathbf{F} . and \mathbf{D} . describe a linear flow:

$$\mathbf{F}_{m+1|m-1} = \mathbf{F}_{m+1|m}\mathbf{F}_{m|m-1}, \quad (14)$$

$$\mathbf{D}_{m+1|m-1} = \mathbf{F}_{m+1|m}\mathbf{D}_{m|m-1}\mathbf{F}_{m+1|m} + \mathbf{D}_{m+1|m}. \quad (15)$$

This representation of the prediction density in (6) has the same structure as the expression in [1, sec. IV, (9)] for the filter density. Therefore, the ASD can be calculated following the equations in there. We obtain a Gaussian density

$$p(\mathbf{x}_{k:n}|Z^{k-1}) = \mathcal{N}(\mathbf{x}_{k:n}; \mathbf{x}_{k:n|k-1}, \mathbf{P}_{k:n|k-1}), \quad (16)$$

where the expectation vector $\mathbf{x}_{k:n|k-1}$ and covariance matrix $\mathbf{P}_{k:n|k-1}$ depend on the sequence type of \mathbf{z}_m . In detail it holds:

$$\mathbf{x}_{k:n|k-1} = (\mathbf{x}_{k|k-1}, \dots, \mathbf{x}_{n|k-1})^\top, \quad (17)$$

where

$$\mathbf{x}_{m|k-1} = \begin{cases} \mathbf{F}_{k|k-1}\mathbf{x}_{k-1|k-1}, & \text{if } m = k \\ \mathbf{F}_{m|m-1}\mathbf{x}_{m-1|k-1} + \mathbf{W}_{m|m+1}(\mathbf{x}_{m+1|k-1} - \mathbf{F}_{m+1|m-1}\mathbf{x}_{m-1|k-1}). & \text{if } m < k \end{cases} \quad (18)$$

In the latter case, the predicted state part \mathbf{x}_m is also called *continuous time retrodiction*. See [5] for a detailed discussion. All other states $\mathbf{x}_{l|k-1}$ for $l \neq m$ are not affected in the prediction step. For shortening reasons, we have to refer to [1] for the calculation of $\mathbf{P}_{k:n|k-1}$. In conclusion, one might say that the prediction extends the ASD expectation vector by a new element, which represents the best knowledge we have about the state \mathbf{x}_m a-priori. This is equal to the standard Kalman prediction, if the measurement is in-Sequence. Otherwise, it corresponds to the retrodicted a-priori state.

2.2 ASD filtering update

In this section we describe how to process the measurement \mathbf{z}_m to obtain the density $p(\mathbf{x}_{k:n}|Z^k)$. Here, $Z^k = \{Z^{k-1}, \mathbf{z}_m\}$ holds, because of the renumbering mentioned in the previous section. Let therefore $p(\mathbf{z}_m|\mathbf{x}_m) = \mathcal{N}(\mathbf{z}_m; \mathbf{H}_m\mathbf{x}_m, \mathbf{R}_m)$ be the likelihood function with measure covariance matrix \mathbf{R}_m and measurement matrix \mathbf{H}_m . Using projection matrices $\mathbf{\Pi}_m$ defined by $\mathbf{\Pi}_m\mathbf{x}_{k:n} = \mathbf{x}_m$, where $k \geq m \geq n$, the Kalman Filter equations are valid for the accumulated state $\mathbf{x}_{k:n}$. This is due to the fact that the concatenated projection $\mathbf{\Pi}_m$ and the measurement matrix \mathbf{H}_m become a linear function again. The sensor likelihood function now has the following form:

$$p(\mathbf{z}_m|\mathbf{x}_{k:n}) = \mathcal{N}(\mathbf{z}_m; \mathbf{H}_m\mathbf{\Pi}_m\mathbf{x}_{k:n}, \mathbf{R}_m). \quad (19)$$

Thus, standard reasoning for the Kalman update equations applied to the accumulated state yields by a use of Bayes' law:

$$p(\mathbf{x}_{k:n}|Z^k) \propto p(\mathbf{z}_m|\mathbf{x}_{k:n}) \cdot p(\mathbf{x}_{k:n}|Z^{k-1}) \quad (20)$$

$$= \mathcal{N}(\mathbf{z}_m; \mathbf{H}_m\mathbf{\Pi}_m\mathbf{x}_{k:n}, \mathbf{R}_m) \quad (21)$$

$$\cdot \mathcal{N}(\mathbf{x}_{k:n}; \mathbf{x}_{k:n|k-1}, \mathbf{P}_{k:n|k-1}). \quad (22)$$

Another use of the product formula (51) enables us to derive the desired filtered pdf:

$$p(\mathbf{x}_{k:n}|Z^k) = \mathcal{N}(\mathbf{x}_{k:n}; \mathbf{x}_{k:n|k}, \mathbf{P}_{k:n|k}), \quad (23)$$

where the expectation value $\mathbf{x}_{k:n|k}$ and the covariance matrix $\mathbf{P}_{k:n|k}$ are given by the following expressions:

$$\mathbf{x}_{k:n|k} = \mathbf{x}_{k:n|k-1} + \mathbf{W}_{k:n}(\mathbf{z}_m - \mathbf{H}_m\mathbf{\Pi}_m\mathbf{x}_{k:n|k-1}), \quad (24)$$

$$\mathbf{W}_{k:n} = \mathbf{P}_{k:n|k-1}\mathbf{\Pi}_m^\top\mathbf{H}_m^\top\mathbf{S}_{m|m-1}^{-1}, \quad (25)$$

$$\mathbf{S}_{m|m-1} = \mathbf{H}_m\mathbf{\Pi}_m\mathbf{P}_{k:n|k-1}\mathbf{\Pi}_m^\top\mathbf{H}_m^\top + \mathbf{R}_m, \quad (26)$$

$$\mathbf{P}_{k:n|k} = \mathbf{P}_{k:n|k-1} - \mathbf{W}_{k:n}\mathbf{S}_{m|m-1}\mathbf{W}_{k:n}^\top. \quad (27)$$

Particularly, the above equations allow us to calculate the impact of an OoS measurement with time stamp t_m on every state \mathbf{x}_l with $t_l \in [t_n, t_k]$. To this end, we use the fact that $\mathbf{P}_{k:n|k}$ is symmetric and then pick out the l -th block line. We then have:

$$\mathbf{x}_{l|k} = \mathbf{x}_{l|k-1} + \mathbf{W}_{l|k,m}(\mathbf{z}_m - \mathbf{H}_m\mathbf{x}_{m|k-1}) \quad (28)$$

$$\mathbf{P}_{l|k} = \mathbf{P}_{l|k-1} - \mathbf{W}_{l|k,m}\mathbf{S}_{m|m-1}\mathbf{W}_{l|k,m}^\top, \quad (29)$$

where the accumulated gain matrix $\mathbf{W}_{l|k,m}$ is given by

$$\mathbf{W}_{l|k,m} = \mathbf{W}_{l|m}\mathbf{P}_{\max\{l,m\}|k}\mathbf{W}_{m|l}^\top\mathbf{H}_m^\top\mathbf{S}_{m|m-1}^{-1}. \quad (30)$$

Here, we again define

$$\mathbf{W}_{r|s} = \begin{cases} \mathbf{1} & \text{for all } r \geq s \\ \prod_{\lambda=r}^{s-1} \mathbf{P}_{\lambda|\lambda}\mathbf{F}_{\lambda+1|\lambda}^\top\mathbf{P}_{\lambda+1|\lambda}^{-1} & \text{for } r < s, \end{cases} \quad (31)$$

where $\mathbf{1}$ denotes the identity matrix in the dimension of \mathbf{x}_l . Note that the RTS recursion is already included in the above equations.

3 Extension to MHT/IMM

In many sensor applications, the sensor output can be ambiguous. This ambiguity can arise from imperfect detection and false measurements, which are also known as *clutter*. In this case, the challenge of data association arises. A well-known methodology to overcome this challenge is the MHT [4, 10]. Furthermore, it has been shown that the parallel use of different dynamics models in an IMM in combination with the MHT outperforms most other filter techniques in adequate scenarios [11]. However, the processing of OoS measurements still is an open issue in this case. This section describes an extension of the above concluded ASD tracker to a combined MHT/IMM framework, which is able to handle OoS measurements on the fly.

To this end, assume a sensor with a detection probability P_D and a false return density ρ_F . We further summarize the sensor output at time t_k as $Z_k = \{\mathbf{z}_k^j\}_{j=1, \dots, m_k}$, where m_k denotes the number of measurements at time t_k .

3.1 Multiple Hypothesis Extension

Let $j_k = 0$ denote the data interpretation hypothesis that the object has not been detected at all by the sensor at time t_k , i.e. all sensor data have to be considered as false measurements, while $1 \leq j_k \leq m_k$ represents the hypothesis

that the object has been detected, $\mathbf{z}_k^{j_k} \in Z_k$ being the corresponding measurement of the object properties, the remaining sensor data being false. Evidently, $\{0, \dots, m_k\}$ denotes a set of mutually exclusive and exhaustive data interpretations. Marginalization yields for this example a likelihood function for ambiguous data given by a weighted sum of Gaussians and a constant (see [6], e.g.):

$$p(Z_k|\mathbf{x}_k) = \sum_{j_k=0}^{m_k} p(Z_k|j_k, \mathbf{x}_k) p(j_k) \quad (32)$$

$$\propto (1 - P_D)\rho_F + P_D \sum_{j_k=1}^{m_k} \mathcal{N}(\mathbf{z}_k^{j_k}; \mathbf{H}_k \mathbf{x}_k, \mathbf{R}_k). \quad (33)$$

Data interpretation hypotheses are the basis for MHT techniques (see [7], e.g.). The associated origin of a time series $Z^k = \{Z_k, Z^{k-1}\}$ of sensor data accumulated up to the time t_k can be interpreted by interpretation histories,

$$\mathbf{j}_k = (j_k, \dots, j_1), \quad \text{where } 0 \leq j_l \leq m_l, \quad (34)$$

that assume at each data collection time t_l , $1 \leq l \leq k$, a certain data interpretation j_l to be true. Via marginalization, the filtering density $p(\mathbf{x}_{k:n}|Z^k)$ can be written as a mixture over such interpretation histories \mathbf{j}_k :

$$\begin{aligned} p(\mathbf{x}_{k:n}|Z^k) &= \sum_{\mathbf{j}_k} p(\mathbf{j}_k|Z^k) p(\mathbf{x}_{k:n}|\mathbf{j}_k, Z^k) \\ &=: \sum_{\mathbf{j}_k} p_{k:n|\mathbf{j}_k}^{\mathbf{j}_k} \mathcal{N}(\mathbf{x}_{k:n}; \mathbf{x}_{k:n|\mathbf{j}_k}^{\mathbf{j}_k}, \mathbf{P}_{k:n}^{\mathbf{j}_k}), \end{aligned} \quad (35)$$

where the parameters $\mathbf{x}_{k:n|\mathbf{j}_k}^{\mathbf{j}_k}$ and $\mathbf{P}_{k:n}^{\mathbf{j}_k}$ are calculated by setting evidence to the hypothesis \mathbf{j}_k , i. e. without ambiguity.

3.1.1 Multiple Model Extension

Markovian IMM evolution models (see [12] and the literature cited therein) for object states $\mathbf{x}_k = (\mathbf{x}_k, i_k)$ have the form:

$$\begin{aligned} p(x_k, i_k|x_{k-1}, i_{k-1}) &= \\ p_{i_k i_{k-1}} \mathcal{N}(\mathbf{x}_k; \mathbf{F}_{k|k-1}^{i_k} \mathbf{x}_{k-1}, \mathbf{D}_{k|k-1}^{i_k}). \end{aligned} \quad (36)$$

IMM models are thus characterized by r kinematic linear Gaussian transition densities $p(\mathbf{x}_k|\mathbf{x}_{k-1}, i_k)$ and class transition probabilities $p_{i_k i_{k-1}} = p(i_k|i_{k-1})$ that are to be specified and part of the modeling assumptions.

By making use of the Total Probability Theorem, the IMM approach can be combined with the multiple hypothesis mixture representation of eq. 35:

$$\begin{aligned} p(\mathbf{x}_{k:n}|Z^k) &= \sum_{\mathbf{j}_k} p(\mathbf{j}_k|Z^k) p(\mathbf{x}_{k:n}|\mathbf{j}_k, Z^k) \\ &= \sum_{\mathbf{j}_k} \sum_{i_k, \dots, i_1} p(i_k, \dots, i_1, \mathbf{j}_k|Z^k) \\ &\quad p(\mathbf{x}_{k:n}|i_k, \dots, i_1, \mathbf{j}_k, Z^k). \end{aligned} \quad (37)$$

As the number of mixture components in (37) increases exponentially over time, mixture reduction techniques such as *moment matching* have been developed [9]. Therefore, we now use moment matching in order to approximate the pdf in (37) by a Gaussian mixture with constantly r components for each hypothesis \mathbf{j}_k :

$$\begin{aligned} p(\mathbf{x}_{k:n}|Z^k) &\approx \sum_{\mathbf{j}_k, i_k} p(\mathbf{j}_k, i_k|Z^k) p(\mathbf{x}_{k:n}|\mathbf{j}_k, i_k, Z^k). \\ &= \sum_{\mathbf{j}_k, i_k} p_{k:n|\mathbf{j}_k}^{\mathbf{j}_k i_k} \mathcal{N}(\mathbf{x}_{k:n}; \mathbf{x}_{k:n|\mathbf{j}_k}^{\mathbf{j}_k i_k}, \mathbf{P}_{k:n|\mathbf{j}_k}^{\mathbf{j}_k i_k}) \end{aligned} \quad (38)$$

As we will show below, this form can be achieved by applying each model i_k to the filtering density of the last step. To this end, we state an induction argument. Assume a mixture representation as in (38) for time t_k . In the first part, we calculate the a-priori density (*prediction*), afterwards the a-posteriori density (*filtering*) for the next step $k+1$.

Prediction Let the set of measurements to be processed be $Z_m = \{\mathbf{z}_m^j\}_j$ which was produced at time t_m . As in section 2.1, we can easily achieve a renumbering of the state $\mathbf{x}_{k:n}$ to $\mathbf{x}_{k+1:n}$ such that $t_{k+1} \geq \dots \geq t_n$ and $t_m \in [t_n, t_{k+1}]$.

According to our assumptions, for each hypothesis \mathbf{j}_k there are r mixture components. We keep \mathbf{j}_k fixed and predict each component with every evolution model i_{k+1} . This leads to a sum of r^2 elements:

$$\begin{aligned} p(\mathbf{x}_{k+1:n}|Z^k, \mathbf{j}_k) &= \\ \sum_{i_k, i_{k+1}} p_{k:n|\mathbf{j}_k}^{\mathbf{j}_k i_k} p_{i_{k+1} i_k} \mathcal{N}(\mathbf{x}_{k+1:n}; \mathbf{x}_{k+1:n|\mathbf{j}_k}^{\mathbf{j}_k i_k i_{k+1}}, \mathbf{P}_{k+1:n|\mathbf{j}_k}^{\mathbf{j}_k i_k i_{k+1}}). \end{aligned} \quad (39)$$

The Gaussian densities in this equation can be calculated by evolution model i_{k+1} on the corresponding element for \mathbf{j}_k and i_k of the filtering mixture at time t_k as described in section 2.1. By moment matching, we reduce the number of components to r matched elements. The density of the above equation (39) is then approximated by

$$\begin{aligned} p(\mathbf{x}_{k+1:n}|Z^k, \mathbf{j}_k) &\approx \\ \sum_{i_{k+1}} p_{k+1:n|\mathbf{j}_k}^{\mathbf{j}_k i_{k+1}} \mathcal{N}(\mathbf{x}_{k+1:n}; \mathbf{x}_{k+1:n|\mathbf{j}_k}^{\mathbf{j}_k i_{k+1}}, \mathbf{P}_{k+1:n|\mathbf{j}_k}^{\mathbf{j}_k i_{k+1}}), \end{aligned} \quad (40)$$

where the mentioned parameters are given by:

$$p_{k+1:n|\mathbf{j}_k}^{\mathbf{j}_k i_{k+1}} = \sum_{i_k} p_{k:n|\mathbf{j}_k}^{\mathbf{j}_k i_k} p_{i_{k+1} i_k}, \quad (41)$$

$$\mathbf{x}_{k+1:n|\mathbf{j}_k}^{\mathbf{j}_k i_{k+1}} = (\mathbf{x}_{k+1|k}^{\mathbf{j}_k i_{k+1}}, \dots, \mathbf{x}_{n|k}^{\mathbf{j}_k i_{k+1}})^\top, \quad (42)$$

$$\mathbf{P}_{k+1:n|\mathbf{j}_k}^{\mathbf{j}_k i_{k+1}} = \frac{1}{p_{k+1:n|\mathbf{j}_k}^{\mathbf{j}_k i_{k+1}}} \sum_{i_k} p_{k:n|\mathbf{j}_k}^{\mathbf{j}_k i_k} p_{i_{k+1} i_k} \mathbf{P}_{k+1:n|\mathbf{j}_k}^{\mathbf{j}_k i_k i_{k+1}}, \quad (43)$$

$$\mathbf{P}_{m|k}^{\mathbf{j}_k i_{k+1}} = \frac{1}{p_{k+1:n|k}^{\mathbf{j}_k i_{k+1}}} \sum_{i_k} p_{k:n|k}^{\mathbf{j}_k i_k} p_{i_{k+1} i_k} \cdot \left[\mathbf{\Pi}_m \mathbf{P}_{k+1:n|k}^{\mathbf{j}_k i_{k+1}} \mathbf{\Pi}_m^\top + (\mathbf{\Pi}_m \mathbf{x}_{k+1:n|k}^{\mathbf{j}_k i_{k+1}} - \mathbf{x}_{m|k}^{\mathbf{j}_k i_{k+1}}) (\mathbf{\Pi}_m \mathbf{x}_{k+1:n|k}^{\mathbf{j}_k i_{k+1}} - \mathbf{x}_{m|k}^{\mathbf{j}_k i_{k+1}})^\top \right]. \quad (44)$$

Mixing up all hypothesis we directly obtain the desired prediction density:

$$p(\mathbf{x}_{k+1:n}|Z^k) \approx \sum_{\mathbf{j}_k, i_{k+1}} p_{k+1:n|k}^{\mathbf{j}_k i_{k+1}} \mathcal{N}(\mathbf{x}_{k+1:n}; \mathbf{x}_{k+1:n|k}^{\mathbf{j}_k i_{k+1}}, \mathbf{P}_{k+1:n|k}^{\mathbf{j}_k i_{k+1}}). \quad (45)$$

Filtering In order to calculate the a-posteriori density we continue a given hypothesis \mathbf{j}_k to $\mathbf{j}_{k+1} = (j_{k+1}, \mathbf{j}_k)$, where j_{k+1} indicates the measurement $\mathbf{z}_m^{j_{k+1}}$ to origin from the observed object. As previously mentioned, $j_{k+1} = 0$ indicates a continuation of the track without using any of the measurements of Z_m at all.

By a use of Bayes law we get:

$$p(\mathbf{x}_{k+1:n}|Z^{k+1}) = \frac{p(Z_m|\mathbf{x}_{k+1:n}) p(\mathbf{x}_{k+1:n}|Z^k)}{\int d\mathbf{x}_{k+1:n} p(Z_m|\mathbf{x}_{k+1:n}) p(\mathbf{x}_{k+1:n}|Z^k)}, \quad (46)$$

because of the renumbering mentioned in the previous section: $Z^{k+1} = Z^k \cup Z_m$. Now we use the sensor likelihood function of (32) and the a-priori density of the previous section. A use of the product formula (51) then yields:

$$p(\mathbf{x}_{k+1:n}|Z^{k+1}) \propto \sum_{\mathbf{j}_{k+1}} p_{k+1:n|k+1}^{\mathbf{j}_{k+1} i_{k+1}} \mathcal{N}(\mathbf{x}_{k+1:n}; \mathbf{x}_{k+1:n|k+1}^{\mathbf{j}_{k+1} i_{k+1}}, \mathbf{P}_{k+1:n|k+1}^{\mathbf{j}_{k+1} i_{k+1}}), \quad (47)$$

with the following abbreviations:

$$p_{k+1:n|k+1}^{\mathbf{j}_{k+1} i_{k+1}} = \frac{\hat{p}_{k+1:n|k+1}^{\mathbf{j}_{k+1} i_{k+1}}}{\sum \hat{p}_{k+1:n|k+1}^{\mathbf{j}_{k+1} i_{k+1}}} \quad (48)$$

$$\hat{p}_{k+1:n|k+1}^{\mathbf{j}_{k+1} i_{k+1}} = p_{k+1:n|k}^{\mathbf{j}_k i_{k+1}} \cdot \begin{cases} (1 - P_D) \rho_F, & \text{if } j_{k+1} = 0 \\ P_D \mathcal{N}(\mathbf{z}_m^{j_{k+1}} - \mathbf{H}_m \mathbf{\Pi}_m \mathbf{x}_{k+1:n|k+1}^{\mathbf{j}_{k+1} i_{k+1}}, \mathbf{S}_{m|m-1}^{\mathbf{j}_{k+1} i_{k+1}}) \end{cases} \quad (49)$$

$$\mathbf{S}_{m|m-1}^{\mathbf{j}_{k+1} i_{k+1}} = \mathbf{H}_m \mathbf{\Pi}_m \mathbf{P}_{k+1:n|k}^{\mathbf{j}_k i_{k+1}} \mathbf{\Pi}_m^\top \mathbf{H}_m^\top + \mathbf{R}_m \quad (50)$$

and the parameters $\mathbf{x}_{k+1:n|k+1}^{\mathbf{j}_{k+1} i_{k+1}}$ and $\mathbf{P}_{k+1:n|k+1}^{\mathbf{j}_{k+1} i_{k+1}}$ given by the ASD update with an application of $\mathbf{z}_m^{j_{k+1}}$ as in section 2.2.

Mixture reduction techniques, such as described in [9], keep the number of mixture components involved manageable. This is done by pruning a hypothesis, if its weight is below a certain threshold. Furthermore, similar hypotheses can be merged by moment matching.

4 Discussion and Results

This section considers aspects of the accuracy achievable by the presented algorithm. Due to necessary approximations as e.g. in (40), the gained final suffers certain limitations. This is discussed in detail in the next section. The results shown in the section afterwards evaluate the ability of the filter to cope with clutter, OoS measurements, and multiple system models.

4.1 Discussion

An important difference to the standard MHT algorithm is that each weight now belongs to all states in the time window $k : n$ of one mixture component. This is due to the fact that each measurement affects every state within the time window. Ambiguous measurements automatically create multiple hypothesis with distinct state quantities at each time step. Furthermore, one should note that we do not recommend the use of dynamics histories i_n, \dots, i_k in order to differentiate the models in use for each state within the time window. The paradigm presented in this paper proposes one mixture component for each system model and hypothesis. This is due to the exponential growth of the number of mixture components. A reasonable example of $r = 3$ system models and a time window of length $k - n = 20$ already yields more than a billion mixture elements for each hypothesis. Therefore, the computational costs would exceed the gain in approximation accuracy.

4.2 Evaluation

In the chosen simulation scenario, an observed object moves through a virtual environment. A sensor measures the position of the object once a second with a normal distributed error $v \sim \mathcal{N}(0, \mathbf{R})$ where $R = 30m \cdot \mathbf{1}$. For each scan, 10 uniformly distributed clutter measurements are added. These measurement sets are filled up into a communication queue where the time stamp of each is obfuscated with a normal distributed delay $\Delta t \sim \mathcal{N}(0s, 10)$. This leads to OoS measurements, due to swapped arrival times.

For the prediction step, two dynamic models are in use. The first model, called *static model*, uses the identity as evolution transition matrix $\mathbf{F}_{k|k-1}$. The second model, called *dynamic model* also respects velocity (and acceleration) by the Newton laws for dynamics. The covariance matrices are the same for both models. The first one should be in use when the object does not move, while the second one matches activity more precisely. In order to evaluate a model switching effect on the filter parameters, the simulation is configured such that the object waits 30 steps, then moves with a constant velocity of $8 \frac{m}{s}$ until step 350, and doesn't move anymore until the end.

Results In figure 1 the effective model for each step in a simulation run can be seen. A point at model=1 represents the static model being in effect at this step, while model=2 indicates the dynamic model to be in charge. Every point inbetween is a real mixture of both with the available

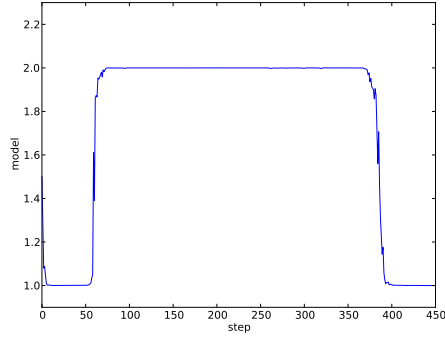


Figure 1: Effective model in use for each step

weight factors being spread over the two models. The curve matches the status of the object adequate, up to a switching delay of a few steps. This delay is enhanced by the OoS measurements of the sensor. In conclusion, one might say that the model switching of the filter is not affected by the OoS measurements.

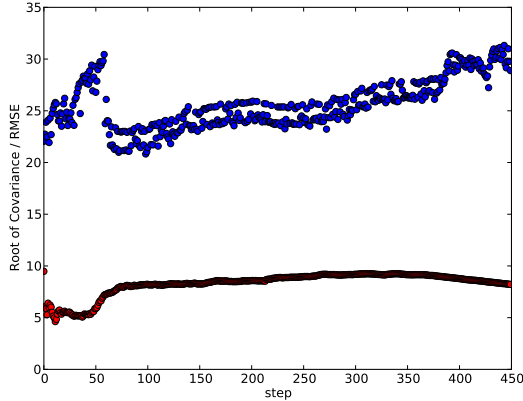


Figure 2: Covariance values and Root Mean Squared Error (RMSE)

In figure 2, one can see the square root of the filter covariance trace $tr(\mathbf{P}_{m|k}^{j_k})$ (as blue dots) and the root mean squared error (RMSE) (red dots). After a few steps, the RMSE levels off at about $8m$. The filter covariance of the single states matches the RMSE quite well, because it is not underestimated. Although the filter is initialized with one hypothesis per measurement in the first set, there was hardly any step with more than one hypothesis occurrence. As a consequence, one can say that the filter easily copes with OoS measurements although there was a lot of clutter present.

Appendix

Product formula

For matrices of suitable dimensions the following formula for products of Gaussians holds:

$$\mathcal{N}(\mathbf{z}; \mathbf{H}\mathbf{x}, \mathbf{R}) \mathcal{N}(\mathbf{x}; \mathbf{y}, \mathbf{P}) = \mathcal{N}(\mathbf{z}; \mathbf{H}\mathbf{y}, \mathbf{S}) \begin{cases} \mathcal{N}(\mathbf{x}; \mathbf{y} + \mathbf{W}\boldsymbol{\nu}, \mathbf{P} - \mathbf{W}\mathbf{S}\mathbf{W}^\top) \\ \mathcal{N}(\mathbf{x}; \mathbf{Q}(\mathbf{P}^{-1}\mathbf{y} + \mathbf{H}^\top\mathbf{R}^{-1}\mathbf{z}), \mathbf{Q}) \end{cases} \quad (51)$$

with the following abbreviations:

$$\boldsymbol{\nu} = \mathbf{z} - \mathbf{H}\mathbf{y}, \quad (52)$$

$$\mathbf{S} = \mathbf{H}\mathbf{P}\mathbf{H}^\top + \mathbf{R}, \quad (53)$$

$$\mathbf{W} = \mathbf{P}\mathbf{H}^\top\mathbf{S}^{-1}, \quad (54)$$

$$\mathbf{Q}^{-1} = \mathbf{P}^{-1} + \mathbf{H}^\top\mathbf{R}^{-1}\mathbf{H}. \quad (55)$$

Sketch of a proof: Interpret $\mathcal{N}(\mathbf{z}; \mathbf{H}\mathbf{x}, \mathbf{R}) \mathcal{N}(\mathbf{x}; \mathbf{y}, \mathbf{P})$ as a joint density $p(\mathbf{z}, \mathbf{x}) = p(\mathbf{z}|\mathbf{x})p(\mathbf{x})$. It can be written as a Gaussian, from which the marginal and conditional densities $p(\mathbf{z})$, $p(\mathbf{x}|\mathbf{z})$ can be derived. In the calculations we make use of known formulae for the inverse of a partitioned matrix (see [3, p. 22], e.g.). From $p(\mathbf{z}, \mathbf{x}) = p(\mathbf{x}|\mathbf{z})p(\mathbf{z})$ the formula results.

References

- [1] W. Koch. "On Accumulated State Densities with Applications to Out-of-Sequence Measurement Processing". In Proc. of the *12th ISIF International Conference on Information Fusion, FUSION 2009*, Seattle, USA, July 6-9, 2009.
- [2] Y. Bar Shalom. "Update with Out-of-Sequence Measurements in Tracking: Exact Solution". *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 38, No. 3, p. 769-777, 2002.
- [3] Y. Bar-Shalom, X.-R. Li, and T. Kirubarajan. "Estimation with Applications to Tracking and Navigation". *Wiley & Sons*, 2001.
- [4] S. Blackman and R. Popoli. "Design and Analysis of Modern Tracking Systems". *Artech House*, 1999.
- [5] W. Koch. "Fixed-Interval Retrodiction Approach to Bayesian IMM-MHT for Maneuvering Multiple Targets". *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 36, No. 1, p. 2-14, 2002.
- [6] W. Koch. "Target Tracking". Chapter 8 in *S. Stergiopoulos (Ed.), Signal Processing Handbook*. CRC Press, 2001.
- [7] S. Blackman. "Multiple Hypothesis Methods". *IEEE Aerospace and Electronic Systems Society Magazine*, **19**, 1, p. 41-52, 2004.

- [8] A. Gelb (Ed.). “Applied Optimal Estimation”, *MIT Press* (1974).
- [9] D. Salmond. “Mixture Reduction Algorithms for Target Tracking in Clutter”. In *SPIE Signal and Data Processing of Small Targets*, vol. 1305, pp. 434-445, (1990).
- [10] D. B. Reid. “An algorithm for tracking multiple targets”. *IEEE Transactions on Automatic Control*, **AC-24**, 6, p. 843–854, 1979.
- [11] S. Blackman, R. Dempster, M. Busch, and R. Popoli. “IMM/MHT solution to radar benchmark tracking problem”. *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 35, No. 2, pp. 730–738, 1999.
- [12] Rong X. Li and V.P. Jilkov. “Survey of maneuvering target tracking. Part V. Multiple model methods”. In *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 41, No. 4, pp. 1255- 1321, 2005.
- [13] Y. Bar-Shalom, H. Chen and M. Mallick, “One-Step Solution for the General Out-of-Sequence Measurement Problem in Tracking”. In *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 40, No. 1, pp. 27–37, 2004

Felix Govaers received his Diploma in Mathematics at the University of Bonn, Germany in 2007. Since then, he has been working at the University of Bonn at the Institute Computer Science 4, department for Communication Systems. Since 2009 he additionally works at the Fraunhofer Institute for Communication, Information and Ergonomics (FKIE), department for Sensor data and Information Fusion (SDF). His research interests are focused on aspects of information fusion in distributed sensor applications.

Wolfgang Koch studied Physics and Mathematics at Aachen Technical University (RWTH), Germany, where he earned a “Dr. rer. nat.” degree in Theoretical Physics. He is head of the dept. “Sensor Data and Information Fusion” at Fraunhofer FKIE, a German defence research lab. At Bonn University, he is teaching Applied Computer Science as a “Privatdozent”. Furthermore he serves the IEEE AESS as an Associate Editor-in-Chief and Technical Editor “Target Tracking and Multisensor Systems” for the IEEE Trans. on AES.