Accurate Likelihood Evaluation for Multiple Model PMHT Algorithms

Tod Luginbuhl
Phillip Ainsleigh
Sunil Mathews
Code 1522
Naval Undersea Warfare Center
Newport, RI 02841-1708
401-832-8241
t.e.luginbuhl@ieee.org

Roy L. Streit
Metron,, Inc.
11911 Freedom Drive
Reston, VA
703-437-2494
streit@metsci.com

Abstract—A variety of authors have incorporated multiple target motion models into the Probabilistic Multi-Hypothesis Tracking (PMHT) algorithm using a discrete Markov chain to model the motion model switching process. However, in these papers the observed data likelihood function is not written down for this model, nor is it evaluated because all possible model assignment sequences must be considered over the PMHT batch. These two issues are addressed in this paper under the assumption that the Markov chain switching model affects the target state process but not the target measurement process: the observed data likelihood function for the PMHT algorithm is given along with a method for evaluating it. A closely related method of including multiple target motion models in the PMHT algorithm that results in a finite mixture distribution of motion models is described as well. In addition, it is shown that using multiple-model smoothing algorithms such as an IMM smoother to estimate the target states in a multiple model PMHT algorithm will not maximize the observed data likelihood function. Finally, it is shown that the MAP target state estimates for linear Gaussian targets with multiple motion models can be computed using a bank of Kalman smoothers. This result fills a gap in the existing literature. [1] [2]

TABLE OF CONTENTS

1. INTRODUCTION

The PMHT algorithm is a multi-target tracking algorithm that estimates target state vectors by maximizing the posterior probability density function (PDF) (i.e. maximum a posteriori or MAP estimates) over a batch of time updates [1, 2]. The target state estimates are obtained using an iterative algorithm derived via the expectation-maximization (EM) algorithm [3, 4]. One of the useful aspects the PMHT algorithm [1, 2] and other methods derived using the EM algorithm is the observed data likelihood function is guaranteed to monotonically ascend during optimization [3, 4]. This is an extremely powerful feature of EM algorithms in general and the PMHT algorithm in particular for at least two reasons. Firstly, good convergence tests can be implemented using the sequence of observed data PDF values, and secondly, because any failure of the observed data PDF to monotonically ascend means there are programming errors or numeric computation problems. As a result, methods for evaluating the observed data likelihood function are highly desirable.

Previous authors incorporated multiple target motion models into the PMHT algorithm to improve tracking of dynamic targets or estimate the time at which a target maneuvers [5-10]. In these references, the target motion models are switched via a discrete Markov chain with known transition matrix. A comparison of the performance of these schemes appears in [6] and [7]. However, the observed data likelihood function (the distribution of the observed data conditioned on the target states) is never explicitly stated or computed in these references. In this paper, the observed data likelihood functions for the multiple-model PMHT algorithm [5-8] and the IMM PMHT algorithm [6, 7] are derived and an exact method for computing them is given in Section 3. In addition, an alternative method of incorporating multiple motion models into the PMHT algorithm is presented in which the discrete target model assignment variables are statistically

[2] IEEEAC paper #1172, Version 4, Updated December 12, 2007

independent. This results in a finite mixture distribution of target motion models rather than a discrete Markov chain, and the evaluation of its observed data likelihood function is straightforward. This approach is referred as the mixed-model PMHT throughout this paper, and the label the multiple-model PMHT is replaced with the jump Markov PMHT to clearly distinguish it from the mixed-model PMHT. This naming convention also more accurately reflects the underlying statistical assumptions.

When multiple target motion models are used in the PMHT algorithm, there is an additional set of missing information, namely, the discrete model assignment variables. The multiple model algorithm can be posed either explicitly by calling out the discrete model assignment variables as missing data inside the PMHT algorithm as in [5-7], or implicitly inside the algorithm used to maximize the auxiliary function. These two algorithms are closely related. In the first algorithm, an iteration first performs data association and target model association and then estimates target states. In the second algorithm, the outer iteration performs data association, and then an inner iteration for each target performs model association followed by state estimation. In principle, the inner iteration of the second algorithm can be performed multiple times for each step of the outer iteration. In this paper, the discrete model assignment variables only affect the evolution of the target state variables and do not affect the measurement distributions (i.e., a target's output distribution is independent of the discrete model switching process). This assumption avoids coupling between targets and, in the authors' experience, the observation distribution is a weak function of the model assignment process that is overwhelmed by the observation distribution's dependence on signal to noise ratio.

For the special case of linear-Gaussian models, it is shown in Section 5 that the EM method leads to the mixed-model and the jump Markov PMHT algorithms that can be implemented via a bank of Kalman smoothers with appropriately modified synthetic system models and input data. This proof fills a gap in the existing literature [17].

2. The PMHT Algorithm

From [1, 2], the probability density function (PDF) of the observed measurements (observed data PDF) equals

$$p(Z,X) = p(Z|X)p(X), \qquad (1)$$

where Z and X represent the collection of measurements and target states over the entire batch respectively. The target state vectors from different targets are statistically independent. In principle, the target state estimates satisfy

$$\hat{X} = \arg\max_X \{p(Z,X)\}; \qquad (2)$$

however, an EM algorithm is only guaranteed to converge to a fixed point of the likelihood function [3, 4]. Therefore, resulting state estimates may correspond to local maxima of (1) rather than the global maximum.

The EM algorithm requires the specification of a set of missing information. In the PMHT algorithm, this set of missing information is the collection of discrete assignment variables that connect the measurements to the targets and clutter distribution. The complete data PDF is the joint distribution of the measurements, the target states, and the discrete assignment variables and is expressed as [1, 2]

$$p(Z,X,K) = p(Z|X,K)p(X)p(K), \qquad (3)$$

where K represents the set of discrete assignment variables. The MAP estimates of the target states are obtained via an auxiliary function that is defined by

$$\mathrm{Q}(X,X') = \sum_K \log\left[p(Z,X,K)\right] p(K|Z,X'), \qquad (4)$$

where

$$p(K|Z,X) = \frac{p(Z,X,K)}{p(Z,X)} \qquad (5)$$

and X' is an initial or previous estimate of the target state vector sequences. [3, 4] prove that maximizing or simply increasing the auxiliary function (4) with respect to the parameters causes the observed data PDF to increase. Because the observed data function is bounded above, alternating between (4) and (5) will cause the observed data PDF (1) to monotonically increase and converge to a stationary point that, for Gaussian densities, is a local maximum.

3. The Observed Data PDF

The complete data probability density function for the PMHT algorithm with multiple target motion models is given by

$$p(Z,X,K,S) = p(Z \mid X,K)p(X \mid S)p(S)p(K), \qquad (6)$$

where Z represents the collection of measurements from the targets and clutter over the batch, X represents the collection of target state vector sequences, K denotes the collection of discrete measurement to target assignment variables, and S denotes the collection of discrete target model assignment variables. The random variables K and S are statistically independent. Under the assumption that all the discrete assignment variables in K are statistically independent, the joint distribution of the measurements and the discrete measurement to target assignment variables conditioned on the collection of target states equals

$$\begin{aligned} p(Z,K \mid X) &= p(Z \mid X,K)p(K) \\ &= \prod_{t=1}^{T}\prod_{n=1}^{N_t} p(z_{tr} \mid x_{tm},k_{tr})\, p(k_{tr} = m) \end{aligned} \tag{7}$$

over the batch of length T where there are $0 \le N_t$ measurements in each scan and $0 \le m \le M$ targets. The clutter distribution is always present and corresponds to $m = 0$, and in this case,

$$p(z_{tr} \mid x_{t0},k_{tr}) = p(z_{tr} \mid k_{tr}) \tag{8}$$

because the clutter distribution has no state process. The joint PDF of the target state variables and the target model assignment variables is expressed as

$$\begin{aligned} p(X,S) &= p(X \mid S)p(S) \\ &= \prod_{m=1}^{M} p(X_m \mid S_m)p(S_m), \end{aligned} \tag{9}$$

where X_m and S_m are the collection of state vectors and discrete target model assignment variables for target m over the batch respectively. The target state vectors evolve according to a continuous Markov process, and the target model assignment variables either are statistically independent (the mixed-model PMHT) or evolve according to a discrete Markov chain (the jump Markov PMHT). For the mixed-model PMHT, the target model assignment variables are independent, hence,

$$\begin{aligned} p(X_m,S_m) &= p(X_m \mid S_m)p(S_m) \\ &= \prod_{t=1}^{T} p(x_{tm} \mid x_{t-1m},s_{tm})p(s_{tm} = j), \end{aligned} \tag{10}$$

where the target model assignment variable, s_{tm}, takes on integer values between one and $J > 1$. For the jump Markov PMHT where there is a Markov chain distribution on the target model assignment, the joint distribution of the target state vectors and the target model assignments equals

$$\begin{aligned} p(X_m \mid S_m)p(S_m) &= p(s_{1m})p(x_{1m} \mid s_{1m}) \\ &\times \prod_{t=2}^{T} p(x_{tm} \mid x_{t-1m},s_{tm})p(s_{tm} \mid s_{t-1m}), \end{aligned} \tag{11}$$

where again s_{tm} takes on integer values between one and $J > 1$.

The observed data PDF conditioned on the target states (i.e., the measurement PDF conditioned only on the target states) is obtained by integrating over all possible discrete measurement to target assignment random variables and discrete target model assignment variables. From [1, 2], the PDF of the measurements conditioned on the target states is

$$\begin{aligned} & p(Z \mid X) = \\ & \prod_{t=1}^{T}\prod_{r=1}^{N_t}\left(\sum_{m=0}^{M} p(z_{tr} \mid x_{tm},k_{tr})\, p(k_{tr} = m)\right). \end{aligned} \tag{12}$$

Unfortunately, no similar expression for the target states is available from the literature. Hence the unconditional PDF of the target states must be derived.

Let $\Psi(T)$ be the set of all discrete target model assignment sequences of length T for all targets. Then, from (9), the probability of all the target states alone is given by

$$p(X) = \sum_{S \in \Psi(T)} p(X \mid S)p(S), \tag{13}$$

where S is a particular set of model assignment sequences for all targets in $\Psi(T)$. Because the targets are statistically independent, the expression in (13) is equivalent to

$$\begin{aligned} p(X) &= \prod_{m=1}^{M} \sum_{S_m \in \Psi_m(T)} p(X_m \mid S_m)p(S_m) \\ &= \prod_{m=1}^{M} p(X_m), \end{aligned} \tag{14}$$

where $\Psi_m(T)$ is the set of all model assignment sequences for target m of length T and S_m is a particular model assignment sequence in $\Psi_m(T)$. Focusing on one particular target, the model assignment sequence can be integrated out directly when the model assignment variables are statistically independent (the mixed-model PMHT), or recursively using the forward part of the forward-backward (Baum-Welch) algorithm [4, 11] when the model assignment variables have a discrete Markov distribution (the jump Markov PMHT). For the mixed-model PMHT, the PDF of a target's state vector sequence equals

$$p\left(X_m\right)=\left(\sum_{j=1}^{J}p\left(x_{1m}\middle|s_{1m}\right)p\left(s_{1m}=j\right)\right)\times\prod_{t=2}^{T}\left(\sum_{j=1}^{J}p\left(x_{tm}\middle|x_{t-1m},s_{tm}\right)p\left(s_{tm}=j\right)\right), \quad (15)$$

which is similar to (12). For the jump Markov PMHT, define the forward probability as

$$\alpha_{mtj}=\sum_{l=1}^{J}p\left(x_{tm}\middle|x_{t-1m},s_{tm}\right)p\left(s_{tm}\middle|l\right)\alpha_{mt-1l}, \quad (16)$$

where the sum over l represents the sum over s_{t-1m} and

$$\alpha_{m1j}=p\left(x_{1m}\middle|s_{1m}\right)p\left(s_{1m}=j\right). \quad (17)$$

Then,

$$p\left(X_m\right)=\sum_{j=1}^{J}\alpha_{mTj} \quad (18)$$

is the PDF of the target's state vector sequence. The observed data PDF for the mixed-model PMHT is obtained by combining equations (12) and (15), or by combining equations (12), (14), and (18) for the jump Markov PMHT. The later case applies to the IMM PMHT because it uses a discrete Markov chain for the model assignment variable as well. Either way, the result

$$p\left(Z,X\right)=p\left(Z\middle|X\right)p\left(X\right), \quad (19)$$

is the same Bayesian form as the observed data PDF (1).

4. TARGET STATE ESTIMATES

The problem of estimating the target state vectors using the EM algorithm can be posed two ways: Only the discrete measurement to target assignment variables are treated as missing information, or the discrete measurement to target assignment variables and the discrete target model assignment variables are treated as missing information. The first case leads to a nested algorithm [12, 13] where the discrete measurement to target assignment variables are treated as missing information in the outer EM algorithm, and the MAP target state estimates are obtained using an inner EM algorithm where the discrete target model assignment variables are treated as missing information. In the second case, both the discrete measurement and model assignment random variables are treated as missing information simultaneously in a single EM algorithm as in [5, 9]. In Section 5, it is shown that MAP target state estimates for linear Gaussian targets with multiple motion models can be computed using a Kalman smoother for each target.

Nested Algorithm

When only the discrete measurement to target assignment variables are treated as missing information and the discrete measurement to target assignment probability mass functions are known, the auxiliary function is given by [1, 2]

$$\mathrm{Q}\left(X,X'\right)=C+\sum_{m=1}^{M}\mathrm{Q}_m\left(X,X'\right), \quad (20)$$

where C is a constant and

$$\mathrm{Q}_m\left(X,X'\right)=\sum_{t=1}^{T}\sum_{r=1}^{N_t}w'_{mtr}\log\left[p\left(z_{tr}\middle|x_{tm},m\right)\right]+\log\left[\sum_{S_m\in\Psi_m(T)}p\left(X_m\middle|S_m\right)p\left(S_m\right)\right]. \quad (21)$$

The probability of a discrete measurement to target assignment variable, k_{tr}, conditioned on measurement, z_{tr}, equals

$$\omega'_{mtr}=p\left(k_{tr}=m\middle|z_{tr},x'_{tm}\right)=\frac{p\left(z_{tr}\middle|x'_{tm},k_{tr}\right)p\left(k_{tr}=m\right)}{\sum_{n=1}^{M}p\left(z_{tr}\middle|x'_{tn},k_{tr}\right)p\left(k_{tr}=n\right)}. \quad (22)$$

In order to maximize the auxiliary function, $Q\left(X,X'\right)$, with respect to the collection of all the target state vector sequences, the maximization step (M-step) of the EM algorithm must maximize each of the functions in (21) with respect to X_m separately. Given that (21) contains the natural logarithm of a sum, the maximization problem appears to be difficult. However, following [1, 2], taking the exponential of (21) yields

$$e^{Q_m\left(X_m\middle|X'_m\right)}=\left[\prod_{t=1}^{T}\prod_{r=1}^{N_t}\left[p\left(z_{tr}\middle|x_{tm},m\right)\right]^{\omega'_{mtr}}\right]\times\left[\sum_{S_m\in\Psi_m(T)}p\left(X_m\middle|S_m\right)p\left(S_m\right)\right], \quad (23)$$

which is equivalent to the original optimization problem because the exponential function is monotonic. Now the optimization problem for each target becomes

$$\hat{X}_m = \arg\max_{X_m} \left\{ \exp\left[Q_m \left(X_m \left| X'_m \right. \right) \right] \right\}. \quad (24)$$

Before discussing how to solve this optimization problem, it is worth making some simplifying assumptions.

The expression in (23) is simplified further under the assumption that

$$p\left(z_{tr} \left| x_{tm}, m \right. \right) = c_{tm} e^{g\left(z_{tr}, x_{tm} \right)}. \quad (25)$$

When this assumption holds, the first term on the right hand side of (23) becomes

$$\exp\left[\sum_{t=1}^{T} \sum_{r=1}^{N_t} \omega'_{mtr} g\left(z_{tr}, x_{tm} \right) \right]. \quad (26)$$

Defining

$$\tilde{g}\left(z_t, x_{tm} \right) = \sum_{r=1}^{N_t} \omega'_{mtr} g\left(z_{tr}, x_{tm} \right), \quad (27)$$

where z_t represents all the measurements in scan t, (23) simplifies to

$$\begin{aligned} p\left(G_m, X_m \right) &= e^{Q_m\left(X_m, X'_m \right)} = \\ &\left[\prod_{t=1}^{T} \tilde{c}_{tm} e^{\tilde{g}\left(z_t, x_{tm} \right)} \right] \\ &\times \left[\sum_{S_m \in \Psi_m(T)} p\left(X_m \left| S_m \right. \right) p\left(S_m \right) \right], \end{aligned} \quad (28)$$

where $\tilde{c}_{tm}$ is an appropriate normalization constant and

$$G_m = \left\{ \tilde{g}\left(z_t, x_{tm} \right) \right\}_{t=1}^{T}. \quad (29)$$

(28) is a fairly general form of the observed data PDF for a jump Markov process where the measurement or output distribution is not switched.

Multiple Model Smoothers—The objective now is to find the sequence of target state vectors that maximize (28) to obtain the MAP estimate. Alternatively, as suggested and explored with the IMM PMHT algorithm [6-8] for linear Gaussian models with Markov chain target model assignments, a multiple model smoothing algorithm [14, 15] can be used to obtain approximate estimates of the target states for the jump Markov PMHT. The IMM PMHT algorithm uses an IMM smoother [14, 15] to compute the target state estimates. More generally, one could defined and use the GPB(n) PMHT algorithm by using a higher order generalized pseudo-Bayesian (GPB) smoother [14, 15] to estimate the target state vectors. However, multiple smoothing algorithms are approximations to the posterior minimum mean square estimator and not the MAP estimator [14-16]. Because (28) is multi-model, the posterior minimum mean square estimates will not equal the MAP estimates, and therefore, using a multiple model smoothing algorithm will not maximize (28). In the early iterations of the PMHT algorithm using a multiple model smoothing algorithm will cause the observed data likelihood function (19) to increase but at some point this monotone ascent will stop because the posterior mean square error estimate does not find the mode of a multi-model posterior distribution. Comparative performance results for linear Gaussian models in [6-8] show that an IMM smoothing algorithm does not perform as well as MAP methods.

MAP Estimates—An EM algorithm can be used to solve the optimization problem in (28) where the discrete target model switching variables are the missing information. To obtain the MAP estimates of the target states using an EM algorithm, the complete data distribution is given by

$$\begin{aligned} &p(G_m, X_m, S_m) = \\ &\tilde{c}_{1m} p\left(x_{1m} \left| s_{1m} \right. \right) p\left(s_{1m} \right) \\ &\times \prod_{t=2}^{T} \tilde{c}_{tm} e^{\tilde{g}\left(z_t, x_{1m} \right)} p\left(x_{tm} \left| x_{t-1m}, s_{tm} \right. \right) p\left(s_{tm} \right), \end{aligned} \quad (30)$$

when the model assignments are independent (from (10) and (28)), or by

$$\begin{aligned} &p\left(G_m, X_m, S_m \right) = \\ &\tilde{c}_{1m} e^{\tilde{g}\left(z_1, x_{1m} \right)} p\left(x_{1m} \left| s_{1m} \right. \right) p\left(s_{1m} \right) \\ &\times \prod_{t=2}^{T} \tilde{c}_{tm} e^{\tilde{g}\left(z_t, x_{tm} \right)} p\left(x_{tm} \left| x_{t-1m}, s_{tm} \right. \right) p\left(s_{tm} \left| s_{t-1m} \right. \right) \end{aligned} \quad (31)$$

when the model assignment evolve via a Markov chain (from (11) and (28)). Dividing (30) or (31) by (28) and using (15) or (18) for the unconditional target state distribution gives

$$\begin{aligned} p\left(S_m \left| G_m, X_m \right. \right) &= \frac{p\left(G_m, X_m, S_m \right)}{p\left(G_m, X_m \right)} \\ &= \frac{p\left(X_m \left| S_m \right. \right) p\left(S_m \right)}{p\left(X_m \right)} = \prod_{t=1}^{T} \prod_{j=1}^{J} \gamma_{tmj} \end{aligned} \quad (32)$$

as the expression for the posterior distribution of the discrete target model assignment variables conditioned on the measurements and the target's state where the expression in (15) is used for $p\left(X_m \right)$,

$$\gamma_{mtj} = \frac{p\left(x_{tm} \mid x_{t-1m}, s_{tm}\right) p\left(s_{tm} = j\right)}{\sum_{l=1}^{J} p\left(x_{tm} \mid x_{t-1m}, s_{tm}\right) p\left(s_{tm} = l\right)}, \quad (33)$$

for $2 \le t \le T$, and

$$\gamma_{m1j} = \frac{p\left(x_{1m} \middle| s_{1m}\right) p\left(s_{1m} = j\right)}{\sum_{l=1}^{J} p\left(x_{1m} \middle| s_{1m}\right) p\left(s_{1m} = l\right)} \quad (34)$$

for the first time update. These last two expressions follow directly from (11) and (15) for the mixed-model PMHT, but the validity of (33) and (34) for the jump Markov PMHT is less obvious. The forward-backward (Baum-Welch) algorithm [4, 11] provides a straightforward method to evaluate the expressions in the numerators in (33) and (34) for Markov model assignments. Using the definition of α_{tm} from (16) and (17), and defining the backward probability as

$$\beta_{mtj} = \sum_{l=1}^{J} p\left(x_{t+1m} \middle| x_{tm}, l\right) p\left(l \middle| s_{tm}\right) \beta_{mt+1l}, \quad (35)$$

where $\beta_{Tmj} = 1$, then (33) becomes

$$\gamma_{mtj} = \frac{\alpha_{mtj}\beta_{mtj}}{\sum_{l=1}^{J} \alpha_{mtl}\beta_{mtl}}. \quad (36)$$

With this in hand and using (30) through (36), the auxiliary function equals

$$\tilde{Q}_m\left(X_m, X'_m\right) = \tilde{C}_m + \sum_{t=1}^{T} \tilde{g}\left(z_t, x_{tm}\right) + \sum_{j=1}^{J} \gamma'_{mtj} \log\left[p\left(x_{1m} \middle| j\right)\right] + \sum_{t=2}^{T} \sum_{j=1}^{J} \gamma'_{mtj} \log\left[p\left(x_{tm} \middle| x_{t-1m}, j\right)\right] \quad (37)$$

for both the mixed-model and the jump Markov PMHT where $\tilde{C}_m$ contains all the terms independent of the target state vectors. The target state estimates are obtained by maximizing $\tilde{Q}_m$ with respect to X_m which is solved using a dynamic program (Viterbi algorithm). The exact form of the solution depends on the particular distributions on the measurements and the target states. The solution for linear Gaussian targets is given in Section 5.

Joint Algorithm

In this algorithm both the discrete measurement to target assignment variables and the discrete target to model assignment variables are treated as missing information. The complete data likelihood function is given in (6), and the observed data likelihood function equals the product of (12) and (18). Since these two sets of missing information are statistically independent, the PDF of the missing information conditioned on the measurements and a previous estimate of the target states is a product of two terms:

$$\begin{aligned} p\left(K, S \middle| Z, X\right) &= \frac{p\left(Z \mid X, K\right) p\left(K\right) p\left(X \middle| S\right) p\left(S\right)}{p\left(Z \middle| X\right) p\left(X\right)} \\ &= p\left(K \middle| Z, X\right) p\left(S \middle| X\right). \end{aligned} \quad (38)$$

The auxiliary function is then defined as

$$\begin{aligned} Q\left(X, X'\right) = &\sum_{K} \log\left[p\left(Z \middle| X, K\right) p\left(K\right)\right] p\left(K \middle| Z, X'\right) \\ &+ \sum_{S \in \Psi(T)} \log\left[p\left(X \middle| S\right) p\left(S\right)\right] p\left(S \middle| X'\right). \end{aligned} \quad (39)$$

The PDF of the discrete measurement to target assignment variables conditioned on the measurements and *a priori* estimate of the target states equals

$$p\left(K \middle| Z, X\right) = \prod_{t=1}^{T} \prod_{m=1}^{M} \prod_{r=1}^{N_t} \omega_{mtr}, \quad (40)$$

where ω_{mtr} is defined in (22). The PDF of the discrete target model assignment variables is given in (32). The auxiliary function is obtained by substituting (40) and (32) into (39) and simplifying. The resulting auxiliary function has the same form as (20), but in this case

$$\begin{aligned} Q_m\left(X_m, X'_m\right) = &\sum_{j=1}^{J} \gamma'_{m1j} \log\left[p\left(x_{1m} \middle| j\right)\right] + \\ &\sum_{t=2}^{T} \sum_{j=1}^{J} \gamma'_{mtj} \log\left[p\left(x_{tm} \middle| x_{t-1m}, j\right)\right] + \\ &\sum_{t=1}^{T} \sum_{r=1}^{N_t} \omega'_{mtr} \log\left[p\left(z_{rt} \middle| x_{tm}, m\right)\right], \end{aligned} \quad (41)$$

which is the same as the auxiliary function from [5]. Because the auxiliary separates into a sum over functions for each target, the maximization problem can be solved separately for each target. Assuming (25) holds and using the definition from (27), the auxiliary function, (41), for a target simplifies to the expression given in (37) which shows that the nested and joint algorithms are nearly identical. The difference is that, in principle, the nested algorithm iteratively maximizes (37) before updating the data association (i.e. recomputing (22) and (27)).

5. LINEAR GAUSSIAN TARGETS

Linear Gaussian targets are an interesting and commonly used special case that results in a closed form solution for the target state estimates. This estimator turns out to be a Kalman smoother. This is demonstrated in [5-8] and [9] when the target model assignment variables have a discrete Markov chain distribution and only the process noise covariance matrices are switched or a control input is switched respectively. However, it will be shown in this section that a Kalman smoother can obtain the target state estimates under more general conditions for the mixed-model PMHT and the jump Markov PMHT. In particular, the switched process noise model from [5-8] will be shown to be a special case.

For linear Gaussian targets, the measurement PDF for each target equals

$$p\left(z_{tr} \mid x_{tm}, k_{tr}\right) = \mathrm{N}\left(z_{tr}; B_{tm}x_{tm}, R_{tm}\right), \quad (42)$$

where $\mathrm{N}(\bullet)$ represents the Normal density (Gaussian PDF). Hence,

$$g\left(z_{tr}, x_{tm}\right) = -\frac{1}{2}\left(z_{tr} - B_{tr}x_{tm}\right)^t R_{tm}^{-1}\left(z_{tr} - B_{tr}x_{tm}\right). \quad (43)$$

Substituting (43) into (27) gives

$$\tilde{g}\left(z_t, x_{tm}\right) = -\frac{1}{2}\sum_{r=1}^{N_t}\omega'_{mtr}\left(z_{rt} - B_{tm}x_{tm}\right)^t R_{tm}^{-1}\left(z_{rt} - B_{tm}x_{tm}\right). \quad (44)$$

The target state density for each target and each model is given by

$$p\left(x_{tm} \mid x_{t-1m}, s_{tm}\right) = \mathrm{N}\left(x_{tm}; A_{tmj}x_{t-1m}, Q_{tmj}\right)\delta\left[j - s_{tm}\right] \quad (45)$$

and

$$p\left(x_{1m} \mid s_{1m}\right) = \mathrm{N}\left(x_{1m}; \overline{x}_{mj}, P_{0mj}\right)\delta\left[j - s_{1m}\right] \quad (46)$$

where $\delta[\bullet]$ denotes the Kronecker delta function. Taking the natural logarithm of (45) and (46) and substituting the result into (37) along with the expression for $\tilde{g}\left(z_t, x_{tm}\right)$ from (44) and simplifying the measurement term yields

$$\begin{aligned}\tilde{\mathrm{Q}}_m\left(X_m, X'_m\right) = \tilde{C}_m - &\sum_{t=1}^{T}\left(\tilde{z}_{tm} - B_{tm}x_{tm}\right)^t \tilde{R}_{tm}^{-1}\left(\tilde{z}_{tm} - B_{tm}x_{tm}\right) - \\ &\sum_{t=2}^{T}\sum_{j=1}^{J}\gamma'_{tmj}\left(x_{tm} - A_{tmj}x_{t-1m}\right)^t Q_{tmj}^{-1}\left(x_{tm} - A_{tmj}x_{t-1m}\right) - \\ &\sum_{j=1}^{J}\gamma'_{1mj}\left(x_{1m} - \overline{x}_{mj}\right)^t P_{0mj}^{-1}\left(x_{1m} - \overline{x}_{mj}\right),\end{aligned} \quad (47)$$

where the terms that do not involve the target state vectors have been absorbed into $\tilde{C}_m$ and the "synthetic" or perhaps more appropriately the expected measurement and measurement covariance equal

$$\tilde{z}_{tm} = \sum_{r=1}^{N_t}\omega'_{mtr}z_{tr} \quad (48)$$

and

$$\tilde{R}_{tm}^{-1} = \left(\sum_{r=1}^{N_t}\omega'_{mtr}\right)R_{tm}^{-1}, \quad (49)$$

respectively. These last two equations perform the data association and generate the sequence of expected measurements and their covariance matrices for each target. By Theorem 4.1 from [17], (47) is maximized with respect to the target state vectors by the Kalman smoother given in Algorithm II from [17]. Unfortunately, [17] failed to discuss whether or not or prove that the effective process noise information matrices (inverse effective process covariance matrices) are positive definite. The Kalman smoother requires that these matrices be positive definite. This oversight on the part of the authors of [17] is addressed in the following paragraphs.

Theorem 4.1 from [17] states that an equivalent form of (47) is

$$\begin{aligned}\mathrm{Q}_m\left(X_m, X'_m\right) = \widehat{C}_m - &\sum_{t=1}^{T}\left(\tilde{z}_{tm} - B_{tm}x_{tm}\right)^t \tilde{R}_{tm}^{-1}\left(\tilde{z}_{tm} - B_{tm}x_{tm}\right) - \\ &\sum_{t=1}^{T}\left(x_{tm} - \tilde{A}_{tm}x_{t-1m}\right)^t S_{tm}^{-1}\left(x_{tm} - \tilde{A}_{tm}x_{t-1m}\right) - \\ &\left(x_{1m} - \tilde{x}_m\right)^t S_{1m}^{-1}\left(x_{1m} - \tilde{x}_m\right),\end{aligned} \quad (50)$$

where $\widehat{C}_m$ is the sum of all the terms that do not involve the target state vectors,

$$\tilde{x}_m = S_{1m} \sum_{j=1}^{J} \gamma'_{tmj} P_{0mj}^{-1} \bar{x}_{mj}, \tag{51}$$

$$\tilde{P}_{0m}^{-1} = \sum_{j=1}^{J} \gamma'_{tmj} P_{0mj}^{-1}, \tag{52}$$

$$S_{1m}^{-1} = \Sigma_{2m}^{-1} + \tilde{P}_{0m}^{-1}, \tag{53}$$

$$S_{tm}^{-1} = \Sigma_{t+1m}^{-1} + \tilde{Q}_{tm}^{-1}, \tag{54}$$

$$\tilde{Q}_{tm}^{-1} = \sum_{j=1}^{J} \gamma'_{tmj} Q_{tmj}^{-1}, \tag{55}$$

$$\Sigma_{tm}^{-1} = \sum_{j=1}^{J} \gamma'_{tmj} A_{tmj}^{t} Q_{tmj}^{-1} A_{tmj} - \tilde{A}_{tm}^{t} S_{tm}^{-1} \tilde{A}_{tm}, \tag{56}$$

$$\Sigma_{T+1m}^{-1} = 0, \tag{57}$$

and

$$\tilde{A}_{tm} = S_{tm} \sum_{j=1}^{J} \gamma'_{tmj} Q_{tmj}^{-1} A_{tmj}. \tag{58}$$

The matrices $\left\{ S_{tm}^{-1} \right\}$ represent the effective process noise information matrices, and because each Σ_{tm}^{-1} is the difference of two positive definite terms it is not obvious that each effective process noise information matrix is positive definite (this is the problem with Theorem 4.1 from [17]).

To prove that the effective process noise covariance matrices are positive definite, it suffices to prove that Σ_{tm}^{-1} is positive semi-definite for $2 \le t \le T+1$. The matrix Σ_{tm}^{-1} is positive semi-definite if and only if

$$y^t \Sigma_{tm}^{-1} y \ge 0 \tag{59}$$

for all y. Let

$$H_{tm} = \sum_{j=1}^{J} \gamma'_{tmj} A_{tmj}^{t} Q_{tmj}^{-1} A_{tmj}. \tag{60}$$

Because H_{tm} is positive definite for all t, it suffices to show that the maximum value of the constrained optimization problem

$$\max_{y} \left\{ y^t \tilde{A}_{tm}^{t} S_{tm}^{-1} \tilde{A}_{tm} y \right\} \le 1 \tag{61}$$

subject to the constraint

$$y^t H_{tm} y = 1 \tag{62}$$

for all y. To simplify the notation, let

$$y_j = A_{tmj} y \tag{63}$$

and

$$G_j = \gamma'_{tmj} Q_{tmj}. \tag{64}$$

The Lagrangian corresponding to the optimization problem is defined as

$$\mathrm{L}\left(y_1, \ldots, y_J\right) = \left(\sum_{j=1}^{J} G_j y_j \right)^t S_{tm}^{-1} \left(\sum_{j=1}^{J} G_j y_j \right) - \lambda \left(\sum_{j=1}^{J} y_j^t G_j y_j \right). \tag{65}$$

The solution for each y_j is given by

$$y_j = \lambda^{-1} S_{tm} v, \tag{66}$$

where

$$\lambda^2 = v^t S_{tm} \tilde{Q}_{tm}^{-1} S_{tm} v \tag{67}$$

and

$$v = \sum_{j=1}^{J} G_j y_j. \tag{68}$$

The maximum value of (61) is then

$$\frac{v^t S_{tm} \tilde{Q}_{tm}^{-1} S_{tm} \tilde{Q}_{tm}^{-1} S_{tm} v}{v^t S_{tm} \tilde{Q}_{tm}^{-1} S_{tm} v}. \tag{69}$$

Note that (66) requires all y_j and hence, all A_{tmj} to be the same in order for the maximum value to be attained. Let

$$u = \tilde{Q}_{tm}^{-1/2} S_{tm} v \tag{70}$$

where $\tilde{Q}_{tm}^{-1/2}$ is a symmetric square root matrix of $\tilde{Q}_{tm}^{-1}$, then

(69) becomes (using (54)) the Rayleigh quotient

$$\frac{u^t \left[\tilde{Q}_{tm}^{1/2} \Sigma_{t+1m}^{-1} \tilde{Q}_{tm}^{1/2} + I \right]^{-1} u}{u^t u}. \qquad (71)$$

By the Rayleigh principle, (71) achieves its maximum value when u is the eigenvector corresponding to the largest eigenvalue of the matrix in the numerator. If Σ_{t+1m}^{-1} is positive semi-definite, then the maximum value of (71) is less than or equal to one. For $t = T$, this holds because of (57); hence, Σ_{Tm}^{-1} and S_{Tm}^{-1} are positive semi-definite and positive definite, respectively. At $t = T - 1$, since Σ_{Tm}^{-1} is positive semi-definite, the maximum value of (71) is less than or equal to one. Consequently, Σ_{T-1m}^{-1} and S_{T-1m}^{-1} are positive semi-definite and positive definite respectively. The remaining Σ_{tm}^{-1} and S_{tm}^{-1} are positive semi-definite and positive definite by induction.

To relate this result to the process noise inflation model given in [5-8], consider the case where a target's state feedback matrices at each time update are all equal:

$$A_{tmj} = \tilde{A}_{tm} \qquad (72)$$

for all $1 \le j \le J$. This requirement on the state feedback matrices is the same as the process noise inflation model described in [5-8]. This happens because the maximum value of the optimization problem given in (61) and (62) is attained at each time update (this follows from (66)), and therefore, Σ_{tm}^{-1} equals the zero matrix and $S_{tm} = \tilde{Q}_{tm}$ for all t. To demonstrate this, note that (56) simplifies to

$$\Sigma_{tm}^{-1} = \tilde{A}_{tm}^t \left(\tilde{Q}_{tm}^{-1} - \tilde{Q}_{tm}^{-1} S_{tm} \tilde{Q}_{tm}^{-1} \right) \tilde{A}_{tm} \quad (73)$$

for each t. For $t = T$, from (54) $S_{Tm} = \tilde{Q}_{Tm}$ due to (57), so Σ_{Tm}^{-1} equals the zero matrix and $S_{T-1m} = \tilde{Q}_{T-1m}$. Therefore, at time $t = T - 1$, Σ_{T-1m}^{-1} also equals the zero matrix. Now consider some time $2 \le t < T - 1$ and assume Σ_{t+1m}^{-1} equals the zero matrix, then $S_{tm} = \tilde{Q}_{tm}$ and by (73) Σ_{tm}^{-1} equals the zero matrix. Hence, by induction, Σ_{tm}^{-1} equals the zero matrix and $S_{tm} = \tilde{Q}_{tm}$ for all $2 \le t \le T + 1$.

6. SUMMARY

In this paper, the problem of incorporating multiple target models into the PMHT algorithm has been discussed. The discrete random variables that determine which target motion model is in use at each time update are modeled by a discrete Markov process or treated as statistically independent and results in two variants of the PMHT algorithm called the jump Markov and the mixed-model PMHT. The label jump Markov PMHT is used in place of the previous name multiple-model PMHT to more accurately reflect the statistical assumptions on the discrete model assignment random variable and clearly distinguish the jump Markov PMHT from the mixed-model PMHT. The name mixed-model PMHT reflects the fact that statistically independent discrete model assignment variables result in a finite mixture distribution over the target motion models. The observed data likelihood function and a method for computing it are given for both the mixed-model PMHT, the jump Markov PMHT and IMM PMHT.

Two closely related methods for estimating the target state vector sequences for the jump Markov PMHT and the mixed-model PMHT are presented. The first method is a nested algorithm that is very similar to the original PMHT algorithm in that the associated data for each target is passed to a separate algorithm that computes the MAP target state estimates for each target (a bank of Kalman smoothers for linear Gaussian targets). While multiple model smoothing algorithms from [12, 13] can be used to obtain approximate target state estimates for the jump Markov PMHT (the IMM PMHT is a particular version of this), the observed data likelihood function for the jump Markov PMHT will not be maximized because these multiple model smoothing algorithms minimize the posterior minimum mean square error rather than find the MAP estimate. Hence, a MAP estimation algorithm for the target states is derived for both the jump Markov PMHT and the mixed-model PMHT using an EM algorithm where the discrete target model assignment variables are treated as missing information. The second method for estimating the target state vector sequences modifies the PMHT algorithm so that the discrete measurement to target and the discrete target to motion model assignment variables are treated as missing information jointly as in [5-9]. In this joint approach, data association and target motion model association are performed during each iteration of the algorithm, where as in the nested (first) approach, for each iteration of the PMHT algorithm (data association), multiple iterations of the MAP target state estimator can be performed (target motion model association). Finally, it is shown that for linear Gaussian targets, the MAP target state estimates obtained by the jump Markov PMHT and the mixed-model PMHT can be computed using a Kalman smoother.

ACKNOWLEDGEMENTS

This work was funded by ONR and OSD.

REFERENCES

[1] R. L. Streit and T. E. Luginbuhl, "Maximum Likelihood Method for Probabilistic Multi-Hypothesis Tracking," in *SPIE Conference on Signal and Data Processing of Small Targets*, Orlando, FL, 1994, pp. 394-405.

[2] R. L. Streit and T. E. Luginbuhl, "Probabilistic Multi-Hypothesis Tracking," Naval Undersea Warfare Center, Newport 10428, February 1995.

[3] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum Likelihood from Incomplete Data via the EM algorithm," *Journal of the Royal Statistical Society Series B*, vol. 39, pp. 1-38, 1977.

[4] L. E. Baum, T. Petrie, G. Soules, and N. Weiss, "A Maximization Technique Occurring in the Statistical Analysis of Probabilistic Functions of Markov Chains," *The Annals of Mathematical Statistics*, vol. 41, pp. 164--171, February 1970.

[5] P. Willett, Y. Ruan, and R. L. Streit, "The PMHT for Maneuvering Targets," in *SPIE Conference on Signal and Data Processing of Small Targets*, Orlando, FL, 1998.

[6] Y. Ruan and P. Willett, "Maneuvering PMHTs," in *SPIE Conference on Signal and Data Processing of Small Targets*, San Diego, CA, 2001, pp. 186-197.

[7] Y. Ruan and P. Willett, "Multiple Model PMHT and its Application to the Second Benchmark Radar Tracking Problem," *IEEE Transactions on Aerospace and Electronic Systems,* vol. 40, pp. 1337-1350, October 2004.

[8] P. Willett, Y. Ruan, and R. L. Streit, "PMHT: Problems and Some Solutions," *IEEE Transactions on Aerospace and Electronic Systems,* vol. 38, pp. 738-754, July 2002.

[9] A. Logothetis, V. Krishnamurthy, and J. Holst, "On Maneuvering Target Tracking via the PMHT," in *Thirty-Sixth Conference on Decision and Control* San Diego, CA, 1997.

[10] G. W. Pulford and B. F. La Scala, "Estimation of Target Manoeuver Sequence with the Expectation-Maximization Algorithm," in *Fourth International Conference on Control, Automation, Robotics, and Vision*, Singapore, 1997, pp. 2340-2344.

[11] L. R. Welch, "Hidden Markov Models and the Baum-Welch Algorithm," *IEEE Information Theory Society Newsletter,* vol. 53, December 2003.

[12] M. Segal and E. Weinstein, "The Cascade EM Algorithm," *Proceedings Of The IEEE,* vol. 76, pp. 1388-1390, October 1988.

[13] D. A. van Dyk, "Nesting EM Algorithms For Computational Efficiency," *Statistica Sinica,* vol. 10, pp. 203-225, 2000.

[14] R. E. Helmick, W. D. Blair, and S. A. Hoffman, "Fixed-Interval Smoothing for Markovian Switching Systems," *IEEE Transactions on Information Theory,* vol. 41, pp. 1845-1855, November 1995.

[15] P. Ainsleigh and T. E. Luginbuhl, "A Family of Multiple-Model Smoothers for Signal Classification," in *2004 IEEE Aerospace Conference*, Big Sky, Montana, 2004.

[16] G. A. Ackerson and K. S. Fu, "On State Estimation in Switching Environments," *IEEE Transactions on Automatic Control,* vol. AC-15, pp. 10-17, February 1970.

[17] A. Logothetis and V. Krishnamurthy, "Expectation Maximization Algorithms for MAP Estimation of Jump Markov Linear Systems," *IEEE Transactions on Signal Processing*, vol. 47, pp. 2139-2156, August 1999.

BIOGRAPHY

Tod Luginbuhl *is a senior researcher in the Sensors and Sonar Systems Department at the Naval Undersea Warfare Center where he has been employed since 1985. He received a B.S. degree in Electrical Engineering from the University of Rochester in 1985. He earned a M.S. and Ph.D. in Electrical Engineering from the University of Connecticut in 1991 and 1999 respectively. His research interests include data association & fusion, frequency tracking, multi-target tracking, probability density function estimation, search theory, sequential analysis, signal classification & detection, and spectrum estimation. He is an associate editor of Sequential Analysis, a senior member of the IEEE and a member Eta Kappa Nu.*

Roy Streit is a Senior Scientist at Metron in Reston, Virginia. His research interests include multi-target tracking, multi-sensor data fusion, and distributed autonomous systems. He has worked at Metron since 2005. Prior to joining Metron, he was a Senior Technologist in the Senior Executive Service at the Naval Undersea Warfare Center in Newport, RI, working primarily on the development, evaluation and application of multi-sensor data fusion algorithms in support of submarine sonar and combat control automation. He has published numerous technical papers, invited and contributed conference papers, and holds nine U.S. patents. He is a Professor (Adjunct) in the Department of Mathematics at the University of Rhode Island and in the Department of Electrical and Computer Engineering at the University of Massachusetts – Dartmouth. He received a Ph.D. in Mathematics in 1978 from the University of Rhode Island. He resides in Rhode Island with his wife, Nancy, and family.

Dr. Phillip Ainsleigh is part of the Sensors and Sonar Systems Department at the Naval Undersea Warfare Center Division Newport. He is currently program manager for a multi-organizational team focused on in-ground fiber-optic (FO) sensor systems for border and perimeter security. As part of that program, Phillip also serves as technical lead on development and implementation of detection, classification, and tracking algorithms for in-ground FO systems. During the period 1998-2006, Phillip worked as principal investigator, individual investigator, and team member while performing basic research in sonar signal classification and tracking. He developed a multi-target track-management system for the interacting-multiple-model/probabilistic-data-association (IMMJPDA) tracking algorithm for passive sonar, as well as signal classification methods employing various types of tracking algorithms. Prior to 1998, Phillip developed signal processing algorithms for acoustic transducer measurements in small pressurized tanks. Phillip received a Ph.D. in electrical engineering from Texas A&M University in 2001. His technical interests focus on signal and information processing algorithms and data fusion methods for a wide range of applications.

Sunil Mathews has been working in the areas of detection and tracking algorithms at the Naval Undersea Warfare Center Division Newport for the past twenty years. His projects have been various from contact association and automated detection and tracking algorithm development, to transitioning of the narrowband contact followers and the (HPMHT) Histogram Probabilistic Multi-hypothesis Tracker to tactical systems. His latest project was adapting the Interacting Multiple Model – Joint Probabilistic Data Association Filter with Track Management (IMM-JPDA/TM) to a tactical array. He has worked as a member of the Advanced Processor Build (APB) tracker community for more than 10 years. Sunil has a Bachelor's in Engineering from the City College of New York and pursuing a Master's in Acoustical Engineering at the Naval Postgraduate School.