Digital Signal Processing
doi:10.1006/dspr.2002.0431

Tracking, Association, and Classification: A Combined PMHT Approach

S. Davey,*,† D. Gray,* and R. Streit‡

*Cooperative Research Center for Sensor Signal and Information processing, Australia, and Electrical and Electronic Engineering Department, The University of Adelaide, South Australia, Australia, †Surveillance Systems Division, DSTO, Australia; and ‡Naval Undersea Warfare Center, Newport, Rhode Island, United States
E-mail: SJDavey@adelaide.edu.au, Doug.Gray@adelaide.edu.au, StreitRL@npt.nuwc.navy.mil,

Davey, S., Gray, D., and Streit, R., Tracking, Association, and Classification: A Combined PMHT Approach, *Digital Signal Processing*.

When tracking more than one object, a key problem is that of associating measurements with particular tracks. Recently, powerful statistical approaches such as probabilistic multihypothesis tracking (PMHT) and probabilistic least squares tracking have been proposed to solve the problem of measurement to track association. However, in practice other information may often be available, typically classification measurements from automatic target recognition algorithms, which help associate certain measurements with particular tracks. An extension to the Bayesian PMHT approach which allows noisy classification measurements to be incorporated in the tracking and association process is derived. Some example results are given to illustrate the performance improvement that can result from this approach.

Key Words: tracking; multitarget association; multihypothesis tracking.

1. INTRODUCTION

The goal of tracking is to produce estimates of the states of targets within a surveillance region. In standard tracking solutions, these states are inferred from a collection of observations each of which may be ambiguous in origin. These observations may be caused by one of several targets present or by undesired processes. Thus the tracking algorithm must associate measurements with state estimates (tracks) and use these measurements to refine the estimates.

The difficult problem in tracking is usually the association of tracks and measurements. This association can be *hard*, where each measurement is

1051-2004/02 $35.00

assigned to a particular track, or *soft*, where the probability of association is used. Algorithms that use *hard* assignments might then attempt to optimize the set of assignments. This is the approach of the multihypothesis tracker (MHT) [1]. The MHT suffers from an exponential growth in the number of assignment hypotheses with time and with the number of targets. A popular algorithm using soft assignments is the probabilistic data association filter (PDAF) [1]. The PDAF enumerates the possible assignments and computes the probability of each. The probabilities are used to generate a synthetic measurement to update the track. The PDAF is a recursive filter but its complexity grows exponentially with the number of targets.

In contrast, the probabilistic multihypothesis tracker (PMHT) [7] treats each assignment as an unknown discrete random variable. Using the expectation maximization (EM) approach [5] the assignment is treated as missing information and the auxiliary function is found by taking the expectation of the likelihood with respect to the assignments. The state estimate is then found by maximizing the auxiliary function. The PMHT has linear complexity in the number of targets and is derived over a batch which provides improved estimation by smoothing.

In some applications, there may be more information available to the tracking algorithm than observations of the target state. In this paper, we consider the case where the algorithm is also provided with an observation of the assignment for each state observation. The assignment observations are referred to as classification measurements. In [1] an algorithm called the *augmented PDAF* is used to incorporate classification information to improve clutter rejection. In this paper, an alternative approach based on the PMHT is presented.

One physical example of classification measurements is high resolution radar. When a target return is distributed through several range bins, the radar receives an image of the target. Features of this image can be used to classify the target [3]. Another example is where data from an ESM receiver may give some classification information on radar targets aligned at the same bearing [2].

This paper is comprised of two sections. First, we show how classification measurements can be incorporated into the PMHT framework. The algorithm to exploit these measurements is then derived. Following the algorithm development, Monte Carlo simulations are used to demonstrate the performance benefits obtainable by using classification measurements. The algorithm presented is appropriate for use in a multitarget tracking environment; however, it is derived for only a single sensor. The standard PMHT has been extended in [4] to use data from multiple dissimilar sensors and this approach could be used to derive a multisensor version of the algorithm presented here.

2. THE PMHT WITH CLASSIFICATION MEASUREMENTS

The derivation of the PMHT with classification measurements follows the same development as the standard PMHT [7]. Let $T \geq 1$ denote the length of the data batch and M denote the assumed number of independent state models. The state of model m at time t is denoted as x_{tm}. We define the sets

$X_t = (x_{t1}, \ldots, x_{tM})$ and $X \equiv (X_0, \ldots, X_T)$. Each model is assumed to have a Markov evolution process $\psi_m\ (x_{tm}|x_{(t-1)m})$ which is known. The probability distribution of the initial value of the state is also assumed to be known for each model and is denoted as $\psi_m\ (x_{0m})$.

The sensor provides a collection of n_t measurements at time instant t, each of which is formed due to a single model. The assignment index k_{tr} denotes the model that caused the rth measurement at time t and takes a value in $[1, M]$. We define the sets $K_t \equiv (k_{t1}, \ldots, k_{tn_t})$ and $K \equiv (K_0, \ldots, K_T)$. Since there are no measurements at $t = 0$, $K_0 = \emptyset$.

The assignments k_{tr} are assumed to be independent of the states x_{tm} and also independent of the assignments of other measurements $k_{\tau\rho}$ with $\tau \neq t$, $\rho \neq r$. Let π_{tm} represent the prior probability that a measurement at time t is associated with model m; i.e., π_{tm} is the probability that $k_{tr} = m$. This probability is assumed to be the same for all measurements at time t; i.e., it does not depend on r. The set $\pi_t \equiv (\pi_{t1}, \ldots, \pi_{tM})$ is the *within scan* measurement probability vector. The set of all assignment probabilities is given by $\Pi \equiv (\pi_1, \ldots, \pi_T)$.

The rth measurement at time t is denoted by z_{tr}. In the standard PMHT, this measurement is an observation of the state of model k_{tr} (k_{tr} is, of course, unknown). In this contribution, the measurement is generalized to contain a continuous observation of the state, $z_{tr}^{(x)}$, and a discrete observation of the assignment index k_{tr} itself, $z_{tr}^{(k)}$; i.e., $z_{tr} = (z_{tr}^{(x)}, z_{tr}^{(k)})$. The discrete component takes integer values in some range $[1, M_k]$ where M_k is the number of possible classes for $z_{tr}^{(k)}$. The sample space of the association measurements may not be the same as the number of targets. An example of a smaller sample space might be if the association measurement was an estimate of target size in the set {small, medium, large}. Define the sets $Z_t \equiv (z_{t1}, \ldots, z_{tn_t})$ and $Z \equiv (Z_1, \ldots, Z_T)$.

For each model, there exists a measurement process $\zeta_m(z_{tr}^{(x)}|x_{tm})$ which is assumed to be known. It is assumed that $z_{tr}^{(k)}$ is independent of the states X. Therefore the probability mass function for the association measurement $P(z_{tr}^{(k)}|k)$ is independent of the state X and the state observation $z_{tr}^{(x)}$. The mass function gives the probability of observing each possible value of $z_{tr}^{(k)}$ for each model and is referred to as the *confusion matrix* in a classification context [8]. The elements of the confusion matrix are denoted as $c_{ij} \equiv P(z_{tr}^{(k)} = i \mid k_{tr} = j)$ with $C = \{c_{ij}\}$. The confusion matrix is assumed to be independent of the measurement index r and time independent. If they are unknown, the c_{ij} are estimated from the observations. If they are known the c_{ij} do not need to be estimated.

As this problem is the same as the usual PMHT problem, except for the inclusion of $z_{tr}^{(k)}$ and c_{ij}, the derivation of the solution is only presented briefly.

The EM algorithm is an iterative solution obtained by maximizing an auxiliary function of the estimates at the previous iteration. The auxiliary function is the expectation of the log likelihood with respect to the missing data, K,

$$Q(X, \Pi, C|X', \Pi', C') = \sum_K \log P(X, Z, K; \Pi, C) P(K|X', Z; \Pi', C'), \qquad (1)$$

where X', Π', and C' indicate estimates from the previous iteration.

As a result of the independence assumptions above, $P(X, Z, K; \Pi, C)$ in (1) may be written as the product of a model term and a measurement term

$$P(X, Z, K; \Pi, C) = P(X) \prod_{t=1}^{T} \left\{ \prod_{r=1}^{n_t} \pi_{t\nu} \zeta_\nu \left(z_{tr}^{(x)} | x_{t\nu} \right) c_{z_{tr}^{(k)}\nu} \Big|_{\nu = k_{tr}} \right\}. \tag{2}$$

The term $P(K|X'Z; \Pi', C')$ in (1) is the probability of the assignments conditioned on the model states and the measurements. Applying Bayes' theorem, this is expressed as $P(K|X', Z; \Pi', C') = P(X', Z, K; \Pi', C') / \sum_K P(X', Z, K; \Pi', C')$. This is then the ratio of (2) to its sum over the associations K. Since the state part of (2) is independent of K it factors out of the sum and cancels with the numerator, giving

$$P(K|X', Z; \Pi', C') = \prod_{t=1}^{T} \prod_{r=1}^{n_t} \frac{\pi_{t\nu} \zeta_\nu (z_{tr}^{(x)} | x_{t\nu}) c_{z_{tr}^{(k)}\nu} |_{\nu = k_{tr}}}{\sum_{k=1}^{M} \pi_{tk} \zeta_k (z_{tr}^{(x)} | x_{tk}) c_{z_{tr}^{(k)} k}} \equiv \prod_{t=1}^{T} \prod_{r=1}^{n_t} w'_{k_{tr} tr}. \tag{3}$$

The difference between the cost function here and that of the standard PMHT [7] is that the association weights $w'_{k_{tr}tr}$ include a term for the association measurement $z_{tr}^{(k)}$ through the confusion matrix. The weight $w'_{k_{tr}tr}$ can be interpreted as the probability that measurement z_{tr} was caused by model k_{tr} given X'_t and the measurement itself. Recall that the measurement consists of both a state observation $z_{tr}^{(x)}$ and an estimate of which model caused this state observation $z_{tr}^{(x)}$.

Equations (2) and (3) are substituted into (1) and after manipulation, the cost function can be written as

$$Q(X, \Pi, C | X', \Pi', C') = \sum_{m=1}^{M} Q_X^m + \sum_{t=1}^{T} Q_{t\Pi} + Q_C. \tag{4}$$

The first term in (4) depends only on the model states and the measurements. This is the same expression that is obtained in the standard PMHT approach except that the w'_{mtr} have the modified form noted earlier. This cost function can be maximized using a Kalman filter when the target measurement and evolution statistics are Gaussian [7]. Similarly, the $Q_{t\Pi}$ term in (4) is the same as that obtained for the standard PMHT but with modified weights. If the prior probabilities Π are known, then this is constant. If they are unknown, they are estimated by maximizing $Q_{t\Pi}$ with respect to Π. This gives the estimator $\hat{\pi}_{tk} = (1/n_t) \sum_{r=1}^{n_t} w'_{ktr}$.

The remaining term in (4), Q_C, is dependent only on the confusion matrix elements c_{ij}. If the elements of the confusion matrix are known, it is constant. If the elements are unknown, estimates of them are found by maximizing Q_C in a similar way to the solution of Π. The Q_C component of the cost function is given by

$$Q_C = \sum_{t=1}^{T} \sum_{r=1}^{n_t} \sum_{k=1}^{M} \log c_{z_{tr}^{(k)} k} w'_{ktr}. \tag{5}$$

This function is to be maximized subject to the constraint that $\sum_i c_{ij} = 1\ \forall j$. An appropriate Lagrangian is therefore $L_C = Q_C + \sum_j \lambda_j\{1 - \sum_i c_{ij}\}$ where λ_j is the Lagrangian multiplier. Differentiating L_C and reapplying the constraint gives

$$\hat{c}_{ij} = \left(\sum_{t=1}^{T} \sum_{r=1}^{n_t} w'_{jtr} \right)^{-1} \sum_{t=1}^{T} \sum_{r=1}^{n_t} w'_{jtr} \delta\left(z_{tr}^{(k)} - i\right), \tag{6}$$

where the identity function $\delta(.)$ takes the value 1 at the origin and zero elsewhere.

This solution assumes a fixed valued confusion matrix. If the confusion matrix is time varying, then the instantaneous estimate will be the same as above with the temporal sums removed. The estimate $\hat{c}_{ij}$ is thus the ratio of the weights for measurements with $z_{tr}^{(k)} = i$ and model j with all the weights for model j. This is an intuitively appealing result.

The algorithm for unknown C and Π can be summarized as follows

1. Initialize $\hat{X}$, $\hat{\Pi}$, and $\hat{C}$.
2. Calculate association weights w_{mtr} for each measurement r and model m at each time sample t defined by $w_{mtr} = \pi_{tm}\zeta_m(z_{tr}^{(x)}|x_{tm})c_{z_{tr}^{(k)}m}$. The weights are normalized by $w_{mtr} = w_{mtr}(\sum_m w_{mtr})^{-1}$. The current estimates $\hat{\pi}_{tm}$ and $\hat{c}_{ij}$ are used in place of π_{tm} and c_{ij}. The current state estimate $\hat{X}$ is used to evaluate ζ_m.
3. Refine the estimate of X using the weights w_{mtr} and the observations z_{tr}. Under linear Gaussian statistics, the solution for X is a fixed interval Kalman smoother using modified measurements defined by $\tilde{z}_t^m = \sum_r w_{mtr} z_{tr}^{(x)}$ with covariance $\tilde{R}_t^m = (\sum_r w_{mtr})^{-1} R^m$ where R^m is the true measurement covariance, i.e., the covariance of the function ζ_m.
4. Refine the estimate of Π using $\hat{\pi}_{tk} = (1/n_t)\sum_r w_{ktr}$.
5. Refine the estimate of C using $\hat{c}_{ij} = (\sum_t \sum_r w_{jtr})^{-1} \sum_t \sum_r w_{jtr}\delta(z_{tr}^{(k)} - i)$.
6. Calculate the cost function Q.
7. Repeat Steps 2–6 until Q converges.

The approach has been to treat the classification measurements as observations of K. An alternative way to incorporate them would be to augment the state vector to include a class for each model and adjust the measurement process accordingly. These approaches are equivalent and the same algorithm results from each.

2.1. Special Cases

We now consider two special cases of this filter. In both cases, we will assume that the PMHT knows the true confusion matrix C.

2.1.1. Uninformative classification measurements. The PMHT with classification measurements is a generalized version of the PMHT. Therefore, we should be able to derive the PMHT algorithm from it.

Suppose that the classification measurements were uninformative, that is each model is equally likely to give rise to each class output. In this case, the confusion matrix is constant across rows. That is, $c_{ij} = c_{ik}\ \forall i, j, k$. Denote c_i as the constant value of the confusion matrix for row i. This situation is

equivalent to having no classification measurements at all since the probability of the observation $z_{tr}^{(k)}$ is independent of the value of k_{tr}. In this case, the weight equations become

$$w'_{k_{tr}tr} = \frac{\pi_{t\nu}\zeta_\nu(z_{tr}^{(x)}|x_{t\nu})c_{z_{tr}^{(k)}}|_{\nu=k_{tr}}}{\sum_{k=1}^{M}\pi_{tk}\zeta_k(z_{tr}^{(x)}|x_{tk})c_{z_{tr}^{(k)}}} = \frac{c_{z_{tr}^{(k)}}\pi_{t\nu}\zeta_\nu(z_{tr}^{(x)}|x_{t\nu})|_{\nu=k_{tr}}}{c_{z_{tr}^{(k)}}\sum_{k=1}^{M}\pi_{tk}\zeta_k(z_{tr}^{(x)}|x_{tk})}.$$

The term due to the classification measurements cancels and the weight equation is the same as that for the standard PMHT. Thus, when the classification measurements are uninformative, the algorithm simplifies to the standard PMHT.

2.1.2. Perfect classification measurements. Another special case of interest is when the output of the classifier is perfect. Suppose we have a separate class for each model and that the classification measurements are error free. In this case, the confusion matrix is the identity matrix and the individual elements may be written as identity functions $c_{ij} = \delta(i - j)$. In this case, the weights become $w'_{k_{tr}tr} = \delta(k_{tr} - z_{tr}^{(k)})$.

This weight equation effectively means that the filter no longer uses probabilistic assignment of the observations and models. This should be expected since the perfect classification measurements are providing the filter with the true values of K. The weights are now hard assignments independent of the model states so there is no need for recursive estimation of the states or the probability parameter Π. The filter simplifies to independent state estimation of the different models using the assignments provided by the classifier as would be expected intuitively.

3. SIMULATION RESULTS

A simple simulation was conducted to test the benefits of using classification measurements with the PMHT. The simulation consisted of two targets (with no clutter) with Gaussian statistics. That is, $\zeta_m(z_{tr}^{(x)}|x_{tm}) \sim N(Hx_{tm}, R)$ and $\psi_m(x_{tm}|x_{(t-1)m}) \sim N(Fx_{(t-1)m}, GQG')$. $N(\mu, S)$ denotes a multivariate normal distribution with mean vector μ and covariance matrix S. The state vector consisted of two components: position and its rate of change. The motion was constant velocity with white Gaussian deviations in the speed. The observations consisted of a measurement of the target position. For this example,

$$F = \begin{bmatrix} 1 & \tau \\ 0 & 1 \end{bmatrix}, \qquad G = \begin{bmatrix} \tau \\ 1 \end{bmatrix}, \qquad H = [1 \quad 0],$$

where τ is the sampling interval, in this case 1. The quantities Q and R are both scalars in this example. To provide almost straight trajectories, Q was set to 0.001. A measurement variance of $R = 1$ was used. At each time sample, a single measurement was taken from either target with equal probability. This corresponds to $n_t = 1$ and $\pi_{t1} = \pi_{t2} = 0.5$ for all t. Since there is only

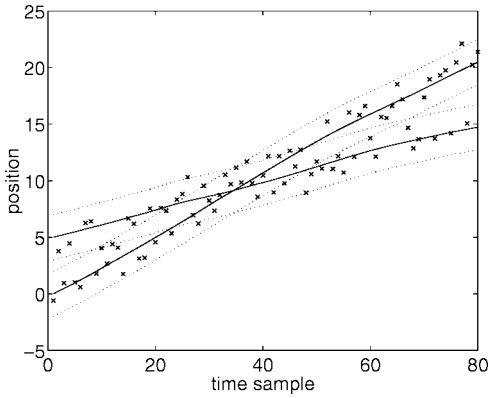

FIG. 1. Example scenario realization.

one measurement per scan, the index r is suppressed in the remainder of the paper. The data were created over a batch of length $T = 80$. Target trajectories were chosen to provide a challenging tracking scenario. The initial states of the targets were $x_{01} = [5, 0.5]^T$ and $x_{02} = [0, 1]^T$. An example realization of the simulated scenario is shown in Fig. 1. The true target trajectories are plotted as solid lines and the observations as crosses. Dotted lines are used to mark curves corresponding to two standard deviations from the mean for the measurement processes ζ_m. These demonstrate the high degree of ambiguity on the $z_t^{(x)}$ measurements.

The classification measurement was a measurement of the source of the state observation and has a confusion matrix of the form

$$C = \begin{bmatrix} \alpha & 1-\alpha \\ 1-\alpha & \alpha \end{bmatrix}.$$

The value of α was varied. By adding the constraint $c_{11} = c_{22}$ to the maximization of Q_C, we obtain an estimator for α:

$$\hat{\alpha} = \frac{1}{T}\left\{ \sum_{t=1}^{T} w'_{1t}\delta\left(z_t^{(k)} - 1\right) + w'_{2t}\delta\left(z_t^{(k)} - 2\right) \right\}. \tag{7}$$

It was assumed that Π was known by the filter. The cases of known and unknown C were both considered. When the confusion matrix was unknown, it was estimated using Eq. (7). A uniform matrix was used as an initial guess for C (i.e., $\alpha = 0.5$). The target state estimates for the filter were initialized by using the correct initial state and projecting a constant velocity path from this point. This scheme is unrealistic in practice, but it avoids irrelevant track initiation issues. The prior state distribution $\psi_m(x_{0m})$ was Gaussian centered on the correct initial state with covariance matrix $P_0 = \mathrm{diag}(R, 0.5)$.

To quantify performance, we examine how well X and C are estimated for various α values. First, we examine how well the state of target 1 is estimated (similar results would be obtained for target 2). The performance is quantified

by plotting the instantaneous RMS state estimation error defined as

$$\epsilon_t = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(\hat{x}_{t1}^n - x_{t1})^T(\hat{x}_{t1}^n - x_{t1})}, \tag{8}$$

where the n superscript denotes the state estimate at the nth random realization and $\hat{x}$ is the state estimate obtained from the PMHT after convergence. This quantity is time dependent because the target geometry changes with time. It is worth noting that the RMS error is not an estimate of the variance of $\hat{x}$ since the estimator is biased. The estimate of the target state will always be biased toward the state of the other target unless the filter has perfect knowledge of the assignments ($\alpha = 1$). The RMS error introduced by bias is, however, small compared with that due to estimated scatter. In all cases, simulations were run over 1000 random realizations.

3.1. RMS Improvement with Known Confusion Matrix

We first investigate the performance gain obtained by using classification measurements when the confusion matrix C is known. The RMS position error is plotted in Fig. 2 for various values of α. As shown in Section 2, the filter is the same as the standard PMHT when α is known to be 0.5. When α is known to be 1, the classifications are perfect and the algorithm simplifies to two independent Kalman smoothers. This case gives the optimal RMS error performance for the given R and Q since the Kalman filter is the optimal filter for the linear Gaussian problem. As is expected, the performance degrades with lower values of α.

The algorithm uses smoothing, which means that measurements from earlier time samples and measurements from later time samples contribute to the state estimate. Because the smoother does not have these extra measurements near the start or the end of the batch, the error rises toward the ends. This is clearly seen in the curve for $\alpha = 1$ which corresponds to simple smoothing. Provided the measurements are good enough, the smoother performance does not depend on

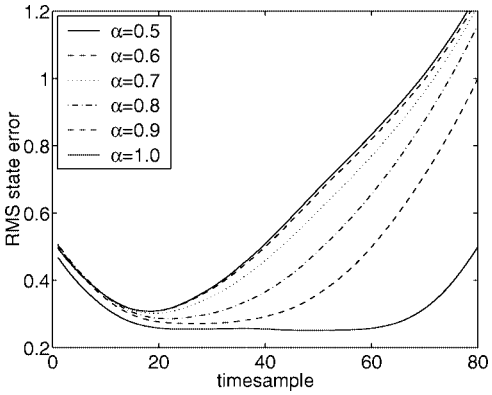

FIG. 2. RMS Position Error, C Known.

the initialization point. However, when the measurements are poor (lower α) the initialization point lowers the error of estimates near it. This is why the minimum error point is near the start of the batch where the initialization still represents the true state well. The dependence on the initialization also explains the gap between the curve for $\alpha = 1$ (which does not depend on initialization) and that for $\alpha < 1$.

As the probability of correct classification increases, the incremental improvement in performance increases at a nonlinear rate. The improvement from $\alpha = 0.8$ to $\alpha = 0.7$ is more than the improvement from $\alpha = 0.7$ to $\alpha = 0.6$.

3.2. RMS Improvement with Unknown Confusion Matrix

We now examine the performance of the PMHT which does not know the confusion matrix C and so estimates α. This was done for various values of α. Figure 3 shows the results when $\alpha = 0.9$. The performance of the filter is plotted with that of the filter that knows C and with the standard PMHT result. As is expected, performance is better when the true confusion matrix is known. However, the approach still shows improvement over the standard PMHT.

Figure 3 shows that, for large values of α, significant performance is lost by estimating the confusion matrix. This is not surprising since we are using a very challenging target scenario. In practice, it may be useful to estimate the confusion matrix during times when the targets are well spaced and then stop adapting the estimate when they become close.

3.3. Confusion Matrix Estimation Performance

Besides the model state, the other quantity estimated by the filter is the probability of correct classification α. Since statistics of the estimator are too difficult to analytically derive, we use simulation to estimate them. We examine the bias and variance of $\hat{\alpha}$ using the sample mean and variance averaged over 1000 Monte Carlo simulations for each true value of α.

In the same way that the model state estimates are biased by the presence of the other model, we expect the α estimates to be biased. One cause of the

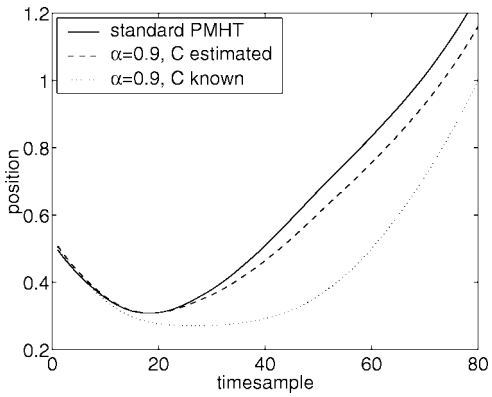

FIG. 3. RMS error, C estimated, $\alpha = 0.9$.

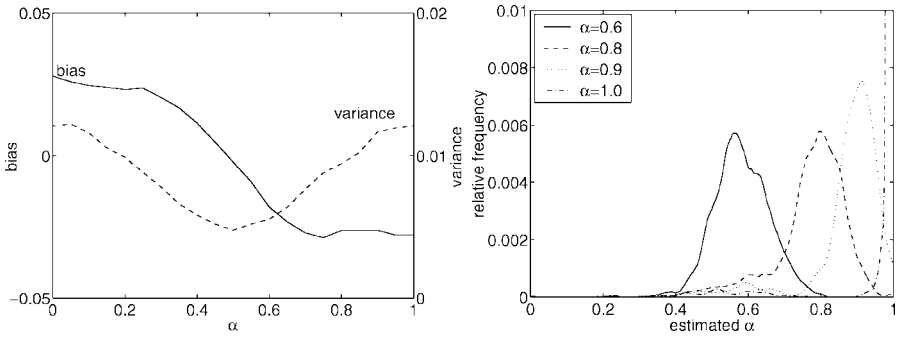

FIG. 4. Bias and variance (left) and estimated probability density of $\hat{\alpha}$ (right).

bias is the hard constraints on $\hat{\alpha}$. These prevent the distribution of $\hat{\alpha}$ from being symmetrical and so produce a bias toward the center of the support of α even if the mode of the distribution is close to α. Also, there occur realizations of the random process which have highly overlapping target paths and state observations that are particularly ambiguous. In these realisations, the filter tends to have trouble estimating the states and also α. The result is often estimates of α close to 0.5.

Figure 4 shows the bias and variance of the estimator $\hat{\alpha}$ as a function of the true α. As we have observed, the estimate is always biased toward 0.5. The bias appears roughly constant at extreme values of α and to transition in a near linear fashion near 0.5. The variance is fairly large and has a minimum at $\alpha = 0.5$. This is somewhat surprising since this implies that the estimation of α becomes more difficult as the classification measurements become more certain. This effect is caused by the troublesome realizations of the target states described above. Figure 4 shows estimated probability density functions for $\hat{\alpha}$ at various values of α. These estimates are obtained by using a kernel density estimate [6]. Clearly, the width of the main peak of the distribution reduces as α increases, as one would expect. Further, it is apparent that the bias is due to the asymmetric tail distribution rather than some systematic error and so it is not anticipated that bias correction would be useful. In a practical implementation of the algorithm, it would be inappropriate to always adapt the confusion matrix estimate. Some target geometries make association very difficult and in such cases the confusion matrix estimate will be poor. Instead, a selective scheme is desirable where the confusion matrix estimate is only updated when the target to measurement association is more certain. This could be automated by choosing measurements that strongly associate with one particular model (i.e., with weights close to zero or unity).

4. CONCLUSION

In some tracking situations, the tracking algorithm has access to more information than ambiguous state observations. In addition to these observations, the algorithm may be provided with measurements that can be treated as observation of the assignment of the state observations. This paper has presented

an extension of the PMHT algorithm that incorporates such information. The benefit of using the additional information has been shown to be significant when the classification measurements are known to be accurate.

REFERENCES

1. Bar-Shalom, Y. and Li, X. R., *Multitarget-Multisensor Tracking: Principles and Techniques.* YBS, Storrs, CT, USA, 1995.
2. Challa, S. and Pulford, G. W., Joint target tracking and classification using radar and ESM sensors. *IEEE Trans. on Aerospace and Electronic Systems* **37** (2001), 1039–1055.
3. Jacobs, S. P. and O'Sullivan, J. A., High resolution radar models for joint tracking and recognition. *Presented at IEEE International Radar Conference, Syracuse, NY, May 1997.*
4. Kreig, M. L. and Gray, D. A., Multi-sensor, probabilistic multi-hypothesis tracking using dissimilar sensors. In *Proceedings of the Conference on Acquisition, Tracking and Pointing XI. SPIE Annual Symposium on Aerosense, Orlando, FL, April 1997.* SPIE, Vol. 3086, pp. 129–138.
5. McLachlan, G. J. and Krishnan, T., *The EM Algorithm and Extensions.* Wiley Interscience, New York, 1997.
6. Silverman, B. W., *Density Estimation.* Chapman and Hall/CRC, New York, 1986.
7. Streit, R. L. and Luginbuhl, T. E., Probabilistic multi-hypothesis tracking. Technical report 10428, Naval Undersea Warfare Center, Feb. 1995.
8. Theodoridis, S. and Koutroumbas, K. *Pattern Recognition.* Academic Press, San Diego, 1999.

SAMUEL J. DAVEY was born in Adelaide, Australia, in 1974. He completed his BE in electrical and electronic engineering at the University of Adelaide in 1995. In 1998 he completed a Master's degree in signal and information processing at the University of Adelaide. He has been employed as a Professional Officer at the Australian Defence Science and Technology Organisation since 1995, where he has worked in the areas of signal processing and automatic tracking. Samuel has been involved in research into advanced automatic tracking systems and performance assessment of automatic tracking systems. He is currently undertaking Ph.D. studies at the University of Adelaide and the Cooperative Research Centre for Sensor Signal and Information Processing on extensions to the probabilistic multiple hypothesis tracking filter.

DOUGLAS A. GRAY holds the Chair of Sensor Signal Processing at Adelaide University, South Australia. He is also Deputy Director of the Cooperative Research Centre for Sensor Signal and Information Processing, leading the Sensor Signal Processing Program. He received the B.Sc(Hons) and Ph.D. in mathematical physics from the University of Adelaide, South Australia, in 1969 and 1973, respectively. From 1973 to 1993 he was with the Australian Defence Science and Technology Organisation, applying signal processing to sonar and electronic warfare. From 1977 to 1979 he was a visiting scientist at the Royal Aircraft Establishment, U.K., and in 1985 was a visiting fellow at the Department of Systems Engineering of the Australian National University. His research interests are in the application of signal processing to sonar, radar, GPS, electronic warfare, and electricity grids, particularly in adaptive processes, beamforming, signal sorting and classification techniques, data fusion, estimation, and system identification.

ROY L. STREIT received the Ph.D. in mathematics from the University of Rhode Island in 1978. He has been with the Naval Undersea Warfare Center since 1970, when he joined one of its predecessor organizations, the Navy Underwater Sound Laboratory in New London, Connecticut. From 1981 to 1982, he was a visiting scholar with the Department of Operations Research, Stanford University. From 1982 to 1984, he was a visiting scientist with the Computer Science Department, Yale University. From 1987 to 1989, he was an echange scientist with the Defence Science and Technology Organisation in Adelaide, Australia. He received the 1999 Solberg Award from the American Society of Naval Engineers for outstanding achievement in research and development related to naval engineering. His current research interests include passive detection and localization of distributed targets in hyperspectral images, tracking in random media using low-fidelity environmental wave propagation models, and numerical acoustic hull array design.