JOURNAL OF GUIDANCE, CONTROL, AND DYNAMICS
Vol. 22, No. 1, January–February 1999

Bayesian Pattern Matching Technique for Target Acquisition

Neil Gordon*
Defence Evaluation and Research Agency, Malvern, Worcester, England WR14 3PS, United Kingdom
David Salmond†
Defence Evaluation and Research Agency, Farnborough, Hants, England GU14 OLX, United Kingdom

The following acquisition/selection problem is considered: a group of N targets is observed at time t_0 and one of them is designated (targeting information). At some later time t_1 the target group is again observed by a missile seeker. The targets are assumed to move as a group, as well as individually, between observation times and so have a dependent motion model. The detection probability at t_1 is less than one, and so some of the targets may not be detected. It is also possible that some measurements received at t_1 may not originate from targets. The problem is to estimate the state of the designated target at time t_1, given the two sets of measurements, i.e., to recover the designated target. We employ a dependent target motion model within a multiple hypothesis framework. The motion of the targets is modeled as the result of two effects: a bulk component, which is common to all targets, and an individual contribution, which is independent from target to target. A closed-form solution is derived for the linear-Gaussian special case, and simulation examples illustrating the technique are presented.

I. Introduction

A KEY problem in missile guidance is that of acquiring the required target in a multiple object scenario. We present a Bayesian technique for assisting this operation. It is assumed that a group of targets is observed at time t_0 via the fire control system, and one of these is designated. This pattern of targets [sometimes known as the target object map (TOM)], including the designation marker, is passed to the missile as targeting information (either prelaunch or via a data link). The missile then flies to the acquisition point and scans the target group with its seeker (at time t_1). The problem is to estimate the state of the designated target given the two sets of measurements, i.e., to recover the designated target from which the acquisition process may be initiated. The algorithm assumes that the targets move as a group, as well as individually, between observation times and that the detection probability at t_1 is less than one, so that some of the targets may not be detected. There is also the possibility that not all measurements received at t_1 originate from targets but may be due to clutter or other effects.

This problem is one of data association. The simplest approach to problems of association uncertainty is the nearest neighbor method: select that association hypothesis that maximizes the measurement likelihood. This type of assignment is usually implemented via an efficient search procedure such as the Munkres algorithm or some variation thereof.[1] An obvious drawback is that the resulting posterior distribution of the target state ignores uncertainty in the chosen measurement association and so is unduly optimistic. Also, if there are a large number of associations having similar likelihood, i.e., if the problem is stressing, there will be a substantial probability of the chosen hypothesis being incorrect. The relationship between the Munkres method and the Bayesian scheme proposed here is discussed more fully in Ref. 2.

The Bayesian multiple hypothesis approach is a better way of modeling significant uncertainty. In this approach, the posterior probabilities of a number of possible associations are evaluated. This allows for the possibility of the most likely option being incorrect and incorporates uncertainty about the association into the posterior distribution. Rather than consider all possible measurement/target associations, it is usual practice to use a gating technique to eliminate very unlikely possibilities. For many interesting problems, a large number of hypotheses may still remain after gating, and it is often desirable to perform a further (posterior) hypothesis reduction after measurement update. This may be necessary to facilitate further processing of the estimator output. The most extreme form of posterior reduction is the probabilistic data association filter (PDAF) approach,[3] where the posterior probability density function (PDF) of the target state is approximated by a single Gaussian. A variety of other reduction methods have been proposed for retention of more information.[1,4]

A common feature of most target tracking/data association algorithms is the assumption that all of the targets present have independent motion models. Exceptions to this are group tracking methods, where the aim is usually to maintain a single track on the group centroid, see Chapter 11 of Ref. 1 and Refs. 5 and 6. In Ref. 5, a general formulation of the problem is given allowing dependence between target motion models; however, an explicit solution including this feature is not attempted. In Ref. 6, the problem is formulated for a Gaussian influence diagram, and a sequence of tests is employed to estimate the association hypothesis. In a discussion of Ref. 6, Speed suggests that it would be better to explicitly model the relationships between targets and, hence, obtain the posterior probabilities of the association and the posterior distribution of the target state. This is the approach adopted for our target selection problem (also see Ref. 7 where an extension of the pattern matching approach to group tracking is described).

In Sec. II we give a detailed problem formulation, and in Sec. III the required probability models are described. A general solution is presented in functional form in Sec. IV. In Sec. V, we consider particular modeling choices that admit an analytic solution. Here all of the models are assumed to be linear and Gaussian, and the posterior PDF is obtained as a Gaussian mixture distribution. In Sec. VI, a hypothesis reduction technique has been developed with the aim of pruning very unlikely hypotheses, which make negligible contribution to the posterior PDF, prior to evaluation. A simulation example illustrating algorithm performance is presented in Sec. VII.

II. Formal Problem Statement

It is assumed that there are N targets present. These are observed by one sensor at time t_0 and then by another sensor (of a possibly different type) at a later time t_1. During the intervening period it is assumed that the number of targets remains constant. The state of target i at time t_k is denoted by $\boldsymbol{x}_{ki}$ where $k = 0, 1$ and $i = 1, \ldots, N$. The target state vector includes position and possibly other attributes. We denote the set of target state vectors at time k by $X_k = (\boldsymbol{x}_{k1}, \ldots, \boldsymbol{x}_{kN})$ where $k = 0$ or 1. The set of measurements obtained at time t_k is denoted by $Z_k = (z_{k1}, \ldots, z_{kN_k})$ where $k = 0$ or 1; these are the two TOMs. At $k = 0$, it is assumed that a clean

Received Sept. 29, 1997; revision received July 20, 1998; accepted for publication July 29, 1998.

*Senior Scientist, Sensors and Processing Sector, EX30, St. Andrews Road. E-mail: gordon@signal.dera.gov.uk.
†DERA Fellow, Weapons Sector, A2 Building, Ively Road.

picture is obtained and every target is detected, so that $N_0 = N$. However, at $k = 1$, each measurement z_{1i} may originate either from one of the targets or it may originate from some other process. Measurements not originating from a target will be generically referred to as spurious measurements. Also at $k = 1$, it is possible that not all of the N targets are detected. For these reasons, in general $N \neq N_1$. Also, the associations between the measurements and the individual targets is uncertain, i.e., z_{ki} does not necessarily originate from $\boldsymbol{x}_{ki}$.

One of the measurements from the set Z_0 is designated. Without loss of generality, the measurements are relabelled so that z_{01} is the designated measurement. It is assumed that the measurement z_{01} originates from a target d, where d is unknown. The problem is to estimate the state $\boldsymbol{x}_{1d}$ of target d at time t_1, after the receipt of the set of measurements Z_1, i.e., we wish to estimate the state of the designated target. For the Bayesian solution it is required to obtain the PDF of $\boldsymbol{x}_{1d}$ given all available information, i.e., $p(\boldsymbol{x}_{1d} \mid Z_1, Z_0)$. This PDF, which will be shown to be a mixture distribution, provides a complete description of the information available on the designated target for acquisition purposes. From it one may derive a point estimate such as the mean or the maximum a posteriori value. Of particular interest for the missile guidance problem is that this PDF precisely quantifies the uncertainty in the identity and state of the designated target, thus providing an indication of the probability of mission success.

Clearly, to solve the preceding problem several relationships have to be modeled. First, it is necessary to model the mechanism by which the sensors extract measurements from the scenario, i.e., the formation of the TOMs. Also, the process by which the scenario evolves with time has to be modeled. To render the problem tractable, it is assumed that only two types of spurious measurement are possible: the first being generated independently of the targets (and subsequently referred to as clutter) and the second type being in some way related to the targets (and so being target state dependent). This second type of spurious measurement might be deliberately or inadvertently deployed by the target. Because there are N targets, the number of possible sources for a spurious measurement is $N+1$. This is composed of N possible sources for spurious measurements caused by targets (assuming all N targets are possible sources of spurious measurements) and one source corresponding to clutter.

The number of targets detected at t_1 is unknown. This is a further layer of uncertainty in addition to the unknown measurement/target association. Also, the unknown association between spurious measurements and their generating sources needs to be considered. In the case of spurious measurements caused by a target, the question of how to deal with the resulting measurement has to be addressed: Is the measurement to be incorporated into the state estimate, i.e., extract information on the target state, or is it simply to be discarded as useless?

A convenient diagrammatic view of the association problem is given in Fig. 1. Starting from the top left of Fig. 1, we see that N measurements contained in Z_0 (the fire control TOM) are extracted from the N target states in X_0. Without loss of generality, the numbering of the targets can be assigned so that z_{0i} originates from $\boldsymbol{x}_{0i}$. The N targets then move according to the bulk and individual motion models to give the states X_1. The second sensor then receives a measurement set Z_1, the seeker TOM. However, we note the following for the measurements in this set.

1) An unknown number of targets N_D is detected where $N_D \leq \min(N, N_1)$.

2) The identity of the N_D detected targets is not known.

3) The generating source of the remaining $(N_1 - N_D)$ measurements is not known.

To model this, mappings or hypotheses λ, τ, and Φ are introduced. The effect of these mappings is to partition the measurement set Z_1, received at time t_1, into a set Z_T (composed of those measurements to be associated as originating from targets) and a set Z_R (composed of those measurements to be associated as spurious measurements). Hence, the measurement association hypothesis may be represented by the compound mapping $\Theta = (\lambda, \Phi, \tau)$. The mapping τ partitions the measurement set Z_1 into Z_T and Z_R. Thus, τ: $\{1, \ldots, N_1\} \rightarrow \{0, 1\}$ with $\tau(i) = 1$ implying that under hypothesis τ, measurement z_{1i} belongs to Z_T and $\tau(i) = 0$ implying that z_{1i} belongs to Z_R. Note that

$$\sum_{i=1}^{N_1} \tau(i) = N_D(\Theta)$$

where $N_D(\Theta)$ is the number of targets detected according to the composite hypothesis Θ. The result of τ is that Z_1 is partitioned into Z_T and Z_R with $Z_T = (z_{T1}, \ldots, z_{TN_D(\Theta)})$ and $Z_R = (z_{R1}, \ldots, z_{R[N_1 - N_D(\Theta)]})$. [Note that Z_T may be uniquely specified from Z_1 and τ according to the rule: for all i, j such that $\tau(i) = \tau(j) = 1$ and $i > j$, if $z_{Tk} = z_{1i}$ and $z_{Tm} = z_{1j}$, then $k > m$, likewise for Z_R.] The mapping λ associates the members of Z_T with targets. It is a mapping from the subscripts of the measurement set Z_T onto the subscripts of the target set X_1. Thus, λ: $\{1, \ldots, N_D(\Theta)\} \rightarrow \{1, \ldots, N\}$ with $\lambda(i) = k$ implying that under hypothesis λ, measurement z_{Ti} is associated with target $\boldsymbol{x}_{1k}$. For convenience, we also define an inverse mapping, ψ: $\{1, \ldots, N\} \rightarrow \{1, \ldots, N_D(\Theta)\}$, such that

$$\psi(i) = \begin{cases} \lambda^{-1}(i) & \text{if} \quad i \in R(\lambda) \\ 0 & \text{otherwise} \end{cases}$$

where $R(\lambda)$ denotes the range of λ, i.e., those target subscripts that are associated with the measurements of Z_1 under hypothesis λ. This is not a true inverse of λ, as the domain of λ has been augmented by 0 to form the range of ψ. The symbols λ and ψ represent the same hypothesis and will be used interchangeably. Finally, Φ associates the members of Z_R with their generating sources, Φ: $\{1, \ldots, N_1 - N_D(\Theta)\} \rightarrow \{0, 1, \ldots, N\}$, where

$$\Phi(i) = \begin{cases} 0 & \text{source of measurement } z_{Ri} \text{ is clutter} \\ m & \text{target } m \text{ is the source of measurement } z_{Ri} \end{cases}$$

with $1 \leq m \leq N$. Note that whereas λ is a one to one mapping, τ and Φ are many to one. For Φ, this is due to the possibility of many spurious measurements being associated with the same source.

It has also been assumed that the prior distributions for each target $\boldsymbol{x}_{0i}$ are identical; hence, because all targets and only targets are observed at time t_0, it is permissible to relabel the targets so that target i is the originator of z_{0i}. Thus, target 1 is the designated target, and the problem is to obtain $p(\boldsymbol{x}_{11} \mid Z_1, Z_0)$.

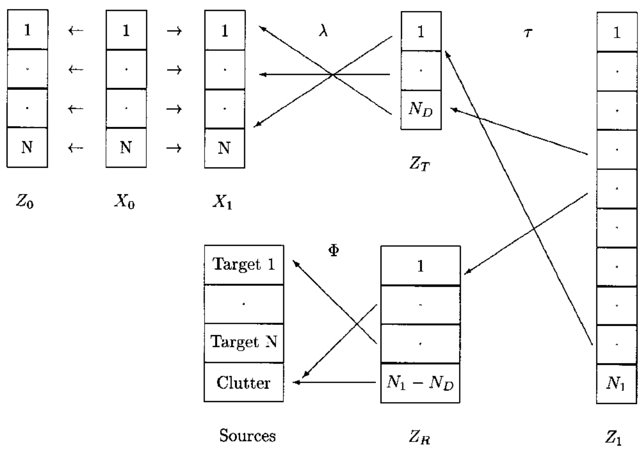

Fig. 1 Measurement associations with spurious measurements.

III. Probability Models

Target Dynamics

The model proposed for the target dynamics and the resulting solution are the most important contributions of this paper. Instead of using the standard assumption that all N targets move independently between time steps, it is assumed that the motion of the targets between t_0 and t_1 is the result of two effects: a bulk component $\boldsymbol{B}$, which is the same for all targets, and an individual component, $\boldsymbol{w}_i$ for target i, which is independent from target to target. Thus, the evolution of the target state may be written

$$\boldsymbol{x}_{1i} = \boldsymbol{f}(\boldsymbol{x}_{0i}, \boldsymbol{B}, \boldsymbol{w}_i) \qquad \text{for} \qquad i = 1, \ldots, N \tag{1}$$

where the function $\boldsymbol{f}$ is the same for all targets. The components $\boldsymbol{B}$ and $\boldsymbol{w}_i$ are unknown although their PDFs $p(\boldsymbol{B})$ and $p(\boldsymbol{w}_i)$ are available. It is assumed that $p(\boldsymbol{w}_i)$ is the same for all targets. The bulk motion $\boldsymbol{B}$ describes the evolution of the target group over the period (t_0, t_1). Note that this bulk component could also be used to model a residual, uncorrected misalignment between the coordinate frames of the two sensors. Equation (1) and $p(\boldsymbol{w}_i)$ define the probabilistic model for the evolution of the target state, $p(\boldsymbol{x}_{1i} \mid \boldsymbol{x}_{0i}, \boldsymbol{B})$. Given $\boldsymbol{B}$, Eq. (1) implies that the target states are independent, and so the target group evolution model is

$$p(X_1 \mid X_0, \boldsymbol{B}) = \prod_{i=1}^{N} p(\boldsymbol{x}_{1i} \mid \boldsymbol{x}_{0i}, \boldsymbol{B}) \tag{2}$$

Measurement Model for Time t_0

The sensor at t_0 is assumed to produce measurements of the form

$$z_{0i} = \boldsymbol{h}_0(\boldsymbol{x}_{0i}, \boldsymbol{v}_{0i}) \tag{3}$$

where measurement i originates from target i. In Eq. (3), $\boldsymbol{v}_{0i}$ is the measurement noise, and its PDF $p(\boldsymbol{v}_{0i})$ is assumed known. Note that $\boldsymbol{v}_{0i}$ is independent of $\boldsymbol{v}_{0j}$ for $j \neq i$. The distribution of $\boldsymbol{v}_{0i}$ is assumed to be the same for all targets. Equation (3) and $p(\boldsymbol{v}_{0i})$ define the probabilistic model for the sensor at time t_0: $p(z_{0i} \mid \boldsymbol{x}_{0i})$. Because the $\boldsymbol{v}_{0i}$ are independent and every target is the origin of exactly one measurement,

$$p(Z_0 \mid X_0) = \prod_{i=1}^{N} p(z_{0i} \mid \boldsymbol{x}_{0i}) \tag{4}$$

Measurement Model for Time t_1

At time t_1, the association between the measurements and their generating sources is unknown. Hence, the measurement model at t_1 is developed conditional on the hypothesis Θ, which links the measurements and their sources. Measurements from the various sources are assumed to be generated as follows:

$$z_{1i} = \begin{cases} \boldsymbol{h}_T(\boldsymbol{x}_{1j}, \boldsymbol{v}_{Ti}) & \text{if target } j \text{ is the source} \\ \boldsymbol{h}_R(\boldsymbol{x}_{1j}, \boldsymbol{v}_{Ri}) & \text{if target } j \text{ is the source of the spurious measurement} \\ \boldsymbol{h}_C(\boldsymbol{v}_{Ci}) & \text{if clutter is the source} \end{cases}$$

where $\boldsymbol{h}_T$, $\boldsymbol{h}_R$, and $\boldsymbol{h}_C$ and the PDFs $p(\boldsymbol{v}_{Ti})$, $p(\boldsymbol{v}_{Ri})$, and $p(\boldsymbol{v}_{Ci})$ are assumed known. Also, $\boldsymbol{v}_{Ti}$, $\boldsymbol{v}_{Ri}$, and $\boldsymbol{v}_{Ci}$ are mutually independent. The distributions of $\boldsymbol{v}_{Ti}$ and $\boldsymbol{v}_{Ri}$ are the same for all targets. Combining spurious measurements from different sources and using the independence of $\boldsymbol{v}_{Ti}$, $\boldsymbol{v}_{Ri}$, and $\boldsymbol{v}_{Ci}$ give the probabilistic measurement model

$$p(Z_1 \mid X_1, N_D, \Theta) = p(Z_T \mid X_1, N_D, \lambda, \tau) p(Z_R \mid X_1, N_D, \Phi, \tau) \tag{5}$$

Model for Number of Targets Detected at t_1

The detection probability P_D is assumed to be constant for each target, and so given that the number of targets is N, the number N_D of targets detected at t_1, has a binomial distribution.

Model for Number of Spurious Measurements at t_1

The number of measurements received at time t_1 is equal to the number of targets detected plus the number of spurious measurements. Thus, in addition to the model for the number of targets detected, a model for the number of spurious measurements is required. This then defines the discrete probability distribution $p(N_1 \mid N_D, N)$, the distribution of the number of measurements received at time t_1 conditional on there being N targets of which N_D are detected. The form assumed for this distribution is clearly problem dependent.

Prior information on Measurement Association at t_1 Given N_1, N, and N_D

Here we make use of the fact that λ is conditionally independent of Φ given τ to obtain

$$p(\Theta \mid N_1, N, N_D) = p(\lambda \mid \tau, N_1, N, N_D) \times p(\Phi \mid \tau, N_1, N, N_D) p(\tau \mid N_1, N, N_D) \tag{6}$$

Because the sensor scan patterns at t_0 and t_1 are assumed independent, for given N_1, N, and N_D, a priori the hypotheses λ are all equally likely, as are the hypotheses τ. Thus, $p(\lambda \mid \tau, N_1, N, N_D)$ may be obtained from the number of possible hypotheses λ given τ, N_1, N, and N_D, and $p(\tau \mid N_1, N, N_D)$ may be obtained from the number of possible hypotheses τ given N_1, N, and N_D (see Ref. 8 for further details). However, the prior probability of hypothesis Φ depends on our prior beliefs concerning the relative chances of each of the available sources generating a spurious measurement. Here we assume that the probability of a spurious measurement being clutter generated is P_C, and all N targets are equally likely to cause spurious measurements. Hence, the probability of a spurious measurement being caused by target i is $(1 - P_C)/N$ for $i = 1, \ldots, N$.

Prior Information on Target States and the Bulk Motion at t_0

It is assumed that the prior information on the target states at time t_0 is independent and the same for each of the N targets, and hence it is given by

$$p(X_0) = \prod_{i=1}^{N} p(\boldsymbol{x}_{0i})$$

This represents very vague information concerning the location of the targets prior to measurements being obtained. Note that no information on the group structure is available prior to the first set of measurements Z_0. The prior distribution of the bulk term $\boldsymbol{B}$ is assumed to be available and is denoted by $p(\boldsymbol{B})$.

IV. General Solution

Our route to obtaining $p(\boldsymbol{x}_{11} \mid Z_1, Z_0)$ is to construct $p(X_1 \mid Z_1, Z_0)$ and then derive the PDF of the required target by integrating out the other members of the target set. It is convenient to write

$$p(X_1 \mid Z_1, Z_0) = \sum_{\Theta} p(X_1 \mid Z_1, Z_0, \Theta) p(\Theta \mid Z_1, Z_0) \tag{7}$$

Here the summation is over every possible hypothesis Θ, $p(X_1 \mid Z_1, Z_0, \Theta)$ is the posterior PDF of the target set X_1 conditional on the compound hypothesis Θ being correct, and $p(\Theta \mid Z_1, Z_0)$ is the posterior probability that Θ is correct. Clearly, to construct this mixture density it is necessary to derive $p(X_1 \mid Z_1, Z_0, \Theta)$ and $p(\Theta \mid Z_1, Z_0)$ and to enumerate every possible hypothesis Θ.

Posterior Probability of X_1 Given Θ

It is convenient to introduce the auxiliary parameter $\boldsymbol{B}$, as the target states are independent given $\boldsymbol{B}$. This allows X_0 and X_1 to be decomposed into individual components

$$p(X_1 \mid Z_1, Z_0, \Theta) = \int p(X_1 \mid Z_1, Z_0, \boldsymbol{B}, \Theta) p(\boldsymbol{B} \mid Z_1, Z_0, \Theta) \, \mathrm{d}\boldsymbol{B} \tag{8}$$

The posterior PDF of X_1 conditional on $\boldsymbol{B}$ and Θ is given by

$$\begin{aligned} p(X_1 \mid Z_1, Z_0, \boldsymbol{B}, \Theta) &\propto p(Z_1 \mid X_1, Z_0, \boldsymbol{B}, \Theta) p(X_1 \mid Z_0, \boldsymbol{B}) \\ &= p(Z_1 \mid X_1, \Theta) \prod_{i=1}^{N} p(\boldsymbol{x}_{1i} \mid z_{0i}, \boldsymbol{B}) \end{aligned} \tag{9}$$

where the likelihood of X_1 can be obtained from the available measurement models. The posterior of $\boldsymbol{B}$ conditional on hypothesis Θ is given by

$$\begin{aligned} p(\boldsymbol{B} \mid Z_1, Z_0, \Theta) &\propto p(\boldsymbol{B}) p(Z_1 \mid Z_0, \boldsymbol{B}, \Theta) \\ &= p(\boldsymbol{B}) \int p(Z_1 \mid X_1, \Theta) p(X_1 \mid Z_0, \boldsymbol{B}) \, \mathrm{d}X_1 \end{aligned} \tag{10}$$

In Eqs. (9) and (10),

$$p(X_1 \mid Z_0, \boldsymbol{B}) = \prod_{i=1}^{N} p(\boldsymbol{x}_{1i} \mid z_{0i}, \boldsymbol{B})$$

the prior of X_1, given Z_0 and $\boldsymbol{B}$, can be obtained from the dynamics model (2), the measurement model (4) for time t_0, and the prior PDF of the target states before any measurement, $p(X_0)$.

Posterior Probability of Θ

To evaluate this probability, the uncertainty in the number of detected targets must be considered. It is convenient to define the notation $Z_1 = (Z_1^*, N_1)$. This emphasizes that Z_1 consists of the dual information of the measurement values Z_1^* and the number N_1 of measurements received at time t_1. Similarly, $Z_0 = (Z_0^*, N)$. Using this notation we have

$$p(\Theta \mid Z_1, Z_0) = p\left(\Theta \,\middle|\, Z_1^*, Z_0^*, N_1, N\right)$$
$$\propto p\left(Z_1^* \,\middle|\, Z_0^*, N_1, N, \Theta\right) p(\Theta \mid N_1, N) \quad (11)$$

where $p(\Theta \mid N_1, N) = p(\Theta \mid Z_0^*, N_1, N)$ because the conditioning on Z_0^* alone contributes no extra information about Θ. The likelihood is simply given by

$$p\left(Z_1^* \,\middle|\, Z_0^*, N_1, N, \Theta\right) = \int p\left(Z_1^* \,\middle|\, Z_0, \boldsymbol{B}, N_1, \Theta\right) p(\boldsymbol{B})\, \mathrm{d}\boldsymbol{B} \quad (12)$$

where $p(Z_1^* \mid Z_0, \boldsymbol{B}, N_1, \Theta)$ can be extracted from Eq. (10). Here, $p(\Theta \mid N_1, N)$ is the probability of the compound hypothesis given N_1 and N. Using the theorem of total probability

$$p(\Theta \mid N_1, N) = \sum_{N_D=0}^{\min(N_1,N)} p(\Theta \mid N_1, N, N_D)\, p(N_D \mid N_1, N) \quad (13)$$

The terms of the summations can be obtained from the prior distribution of Θ given by Eq. (6) and prior models for the numbers of detected targets and spurious measurements at t_1. It can be seen that even to obtain the probability of a single hypothesis, Eq. (11) must be evaluated for every hypothesis to obtain the normalizing denominator. This may entail an enormous computational effort. Also note that the hypothesis probabilities contain other interesting and useful information for missile guidance decisions. For example, the posterior probability that measurement z_{Ti} originated from target k is simply obtained by summing the posterior probability of all hypotheses Θ for which $\lambda(i) = k$. In the same way, the posterior probability that target k was not detected at time t_1 is the sum of $p(\Theta \mid Z_1, Z_0)$ for which $\psi(k) = 0$.

A formal solution for the required PDF, $p(\boldsymbol{x}_{11} \mid Z_1, Z_0)$, may be obtained directly by integrating Eq. (7) over $\boldsymbol{x}_{12}, \boldsymbol{x}_{13}, \ldots, \boldsymbol{x}_{1N}$.

V. Particular Solution for the Linear-Gaussian Problem

In general, it is not possible to obtain a closed-form solution for $p(X_1 \mid Z_1, Z_0)$ due to the integrations involved. In this section we present a special case for which a closed-form solution does exist. The motion and measurement models are all taken to be linear and Gaussian, and all of the target prior distributions are Gaussian.

Following Ref. 3, page 168, the number of spurious measurements is assumed to have either a Poisson distribution or a uniform distribution (Ref. 3 refers to the uniform model as the nonparametric approach). It is further assumed that clutter is the only source of spurious measurements and that they have a uniform spatial distribution. A linear/Gaussian model assumed for the other spurious measurement generating sources would also admit an analytic solution, but for simplicity this case is not shown here. The compound permutation Θ is reduced to only include λ and τ (Φ is redundant). The posterior distribution for X_1 is obtained as a Gaussian mixture. In this case the posterior PDF of any target i (including the required target $i = 1$) is simply obtained as a mixture over the appropriate (Gaussian) marginal PDF of each component of $p(X_1 \mid Z_1, Z_0)$:

$$p(\boldsymbol{x}_{1i} \mid Z_1, Z_0) = \sum_{\Theta} p(\boldsymbol{x}_{1i} \mid Z_1, Z_0, \Theta)\, p(\Theta \mid Z_1, Z_0) \quad (14)$$

The assumed models and prior distributions are taken to be as follows.

1) For target dynamics

$$\boldsymbol{x}_{1i} = \boldsymbol{f}(\boldsymbol{x}_{0i}, \boldsymbol{B}, \boldsymbol{w}_i) = A\boldsymbol{x}_{0i} + \Gamma_1 \boldsymbol{B} + \Gamma_2 \boldsymbol{w}_i$$

where $\boldsymbol{w}_i$ is a Gaussian random variable with zero mean and covariance Q for $i = 1, \ldots, N$. Also, $\boldsymbol{w}_i$ and $\boldsymbol{w}_j$ are assumed independent for $i \neq j$. Thus,

$$p(\boldsymbol{x}_{1i} \mid \boldsymbol{x}_{0i}, \boldsymbol{B}) = N\left(A\boldsymbol{x}_{0i} + \Gamma_1 \boldsymbol{B}, \Gamma_2 Q \Gamma_2^T\right)$$

where $N(\boldsymbol{c}, D)$ is a Gaussian PDF of mean $\boldsymbol{c}$ and covariance D.

2) The measurement equations for the two sensors are of the form

$$\begin{aligned} z_{0i} &= H_0 \boldsymbol{x}_{0i} + \boldsymbol{v}_{0i} && \text{for} \quad i = 1, \ldots, N \\ z_{Ti} &= H_1 \boldsymbol{x}_{1\lambda(i)} + \boldsymbol{v}_{Ti} && \text{for} \quad i = 1, \ldots, N_D \\ z_{Ri} &= \boldsymbol{v}_{Ri} && \text{for} \quad i = 1, \ldots, N_1 - N_D \end{aligned}$$

with $p(\boldsymbol{v}_{0i}) = N(0, R_0)$ for $i = 1, \ldots, N$, $p(\boldsymbol{v}_{Ti}) = N(0, R_1)$ for $i = 1, \ldots, N_D$, and $p(\boldsymbol{v}_{Ri}) = V^{-1}$ for $i = 1, \ldots, N_1 - N_D$, where V is the volume of the search box employed at t_1. Here, $\boldsymbol{v}_{0i}$, $\boldsymbol{v}_{Ti}$, and $\boldsymbol{v}_{Ri}$ are assumed independent. This gives

$$p(z_{0i} \mid \boldsymbol{x}_{0i}) = N(H_0 \boldsymbol{x}_{0i}, R_0)$$
$$p(z_{Ti} \mid \boldsymbol{x}_{1\lambda(i)}, \Theta) = N(H_1 \boldsymbol{x}_{1\lambda(i)}, R_1)$$
$$p(z_{Ri} \mid \Theta) = V^{-1}$$

These models define the likelihood of X_1 in Eq. (9). Note that the association permutation Θ has been included in the conditioning of the measurement model at time t_1.

3) The prior PDF of the target state is taken to be

$$p(X_0) = \prod_{i=1}^{N} p(\boldsymbol{x}_{0i})$$

where $p(\boldsymbol{x}_{0i}) = N(\bar{\boldsymbol{x}}_0, M_0)$ for $i = 1, \ldots, N$. The bulk motion prior is also Gaussian, $p(\boldsymbol{B}) = N(\boldsymbol{b}, \Omega)$.

4) The scan patterns of the sensors at times t_0 and t_1 are assumed to be independent, and so the prior probabilities of the hypotheses are taken to be uniform.

Posterior Distribution of Required Target $\boldsymbol{x}_{11}$ Given Θ: $p(\boldsymbol{x}_{11} \mid Z_1, Z_0, \Theta)$

Using these prior distributions and models, the following distributions may be derived (details are given in Ref. 8).

Posterior Distribution of the Bulk Motion Given Θ: $p(\boldsymbol{B} \mid Z_1, Z_0, \Theta)$

From Eq. (10) $p(\boldsymbol{B} \mid Z_1, Z_0, \Theta) = N[\hat{\boldsymbol{B}}(\Theta), P_B(\Theta)]$, where

$$\hat{\boldsymbol{B}}(\Theta) = \boldsymbol{b} + P_B(\Theta)\Gamma_1^T H_1^T S^{-1} \sum_{i \in R(\lambda)} \boldsymbol{\nu}_i(\Theta) \quad (15)$$

$$P_B(\Theta) = \left[\Omega^{-1} + N_D(\Theta)\Gamma_1^T H_1^T S^{-1} H_1 \Gamma_1\right]^{-1} \quad (16)$$

with

$$S = R_1 + H_1 M_1 H_1^T, \qquad M_1 = A P_0 A^T + \Gamma_2 Q \Gamma_2^T$$
$$P_0 = \left(M_0^{-1} + H_0^T R_0^{-1} H_0\right)^{-1}$$
$$\boldsymbol{\nu}_i(\Theta) = z_{T\psi(i)} - H_1(A\hat{\boldsymbol{x}}_{0i} + \Gamma_1 \boldsymbol{b})$$
$$\hat{\boldsymbol{x}}_{0i} = \bar{\boldsymbol{x}}_0 + K_0(z_{0i} - H_0 \bar{\boldsymbol{x}}_0), \qquad K_0 = P_0 H_0^T R_0^{-1}$$

Note that $\boldsymbol{\nu}_i(\Theta)$ is equivalent to the innovation in a Kalman filter. P_B, $\hat{\boldsymbol{B}}$, and N_D have all been written as a function of Θ to indicate that they depend on the number of associated measurements under Θ.

Posterior Distribution of Target i Given Θ: $p(\boldsymbol{x}_{1i} \mid Z_1, Z_0, \Theta)$

Here,

$$p(\boldsymbol{x}_{1i} \mid Z_1, Z_0, \Theta) = N[\hat{\boldsymbol{x}}_{1i}(\Theta), P_{1i}(\Theta)]$$

where

$$\hat{\boldsymbol{x}}_{1i}(\Theta) = \begin{cases} A\hat{\boldsymbol{x}}_{0i} + \Gamma_1 \hat{\boldsymbol{B}}(\Theta) & \text{if} \quad i \notin R(\lambda) \\ (I - K_1 H_1)\{A\hat{\boldsymbol{x}}_{0i} + \Gamma_1 \hat{\boldsymbol{B}}(\Theta)\} + K_1 z_{T\psi(i)} & \text{if} \quad i \in R(\lambda) \end{cases} \tag{17}$$

and

$$S_\Theta = \begin{pmatrix} S+\Sigma & \Sigma & \cdots & \Sigma \\ \Sigma & S+\Sigma & & \vdots \\ \vdots & & \ddots & \Sigma \\ \Sigma & \cdots & \Sigma & S+\Sigma \end{pmatrix}$$

with

$$\Sigma = H_1 \Gamma_1 \Omega \Gamma_1^T H_1^T$$

$$P_{1i}(\Theta) = \begin{cases} M_1 + \Gamma_1 P_B \Gamma_1^T & \text{if} \quad i \notin R(\lambda) \\ \left(M_1^{-1} + H_1^T R_1^{-1} H_1\right)^{-1} + (I - K_1 H_1)\Gamma_1 P_B \Gamma_1^T (I - K_1 H_1)^T & \text{if} \quad i \in R(\lambda) \end{cases} \tag{18}$$

and

$$K_1 = \left(M_1^{-1} + H_1^T R_1^{-1} H_1\right)^{-1} H_1^T R_1^{-1}$$

Note that there are two results depending on whether or not the hypothesis Θ indicates that target i has been observed. If target i has not been observed, then the predicted position is only updated by the posterior bulk mean. However, if target i has been observed (according to Θ), then this measurement is used to update the predicted position in conjunction with the posterior bulk mean. Note for the designated target set, $i = 1$.

Posterior Probability of Hypothesis Θ: $p(\Theta \mid Z_1, Z_1)$

Solutions are given for both a Poisson distribution and a uniform distribution for the number of spurious measurements received. These clutter models are widely used in target tracking studies.[3]

Number of Spurious Measurements

The distribution of the number of spurious (clutter) measurements is taken to be either Poisson with mean (μV) (where μ is the density of clutter measurements) or uniform. Hence,

$$p(N_1 \mid N_D, N) = \begin{cases} e^{-\mu V}(\mu V)^{N_1 - N_D}/(N_1 - N_D)! & \text{for} \quad N_1 \geq N_D\text{: Poisson model} \\ \delta & \text{for} \quad N_1 \geq N_D\text{: uniform model} \end{cases}$$

where δ is any positive constant.

Posterior Probability of Θ: $p(\Theta \mid Z_1, Z_0)$

Using the prior information and measurement models, the following result may be obtained for $P_D < 1$ and $\mu > 0$ (see Ref. 8):

$$p(\Theta \mid Z_1, Z_0) \propto p(Z_T \mid Z_0, \Theta) p(Z_R \mid \Theta) p(\Theta \mid N_1, N) = \begin{cases} \left[\dfrac{P_D}{\mu(1-P_D)}\right]^{N_D(\Theta)} \times N[\bar{Z}_T(\Theta), S_\Theta] & \text{: Poisson model} \\ [N_1 - N_D(\Theta)] \times \left(\dfrac{V P_D}{1-P_D}\right)^{N_D(\Theta)} \times N[\bar{Z}_T(\Theta), S_\Theta] & \text{: uniform model} \end{cases} \tag{19}$$

with

$$N(\bar{Z}_T(\Theta), S_\Theta) = \frac{\exp(-F(\Theta)/2)}{\sqrt{(2\pi)^{mN_D(\Theta)}|S_\Theta|}}$$

where

$$F(\Theta) = \sum_{i \in R(\lambda)} \boldsymbol{\nu}_i^T(\Theta)\left[S + N_D(\Theta) H_1 \Gamma_1 \Omega \Gamma_1^T H_1^T\right]^{-1} \boldsymbol{\nu}_i(\Theta) + \sum_{i \in R(\lambda)} \sum_{k \in R(\lambda), k > i} [\boldsymbol{\nu}_k(\Theta) - \boldsymbol{\nu}_i(\Theta)]^T \times S^{-1} H_1 \Gamma_1 P_B \Gamma_1^T H_1^T S^{-1} [\boldsymbol{\nu}_k(\Theta) - \boldsymbol{\nu}_i(\Theta)] \tag{20}$$

and where m is the dimension of the measurements z_{1i}. S_Θ contains $N_D(\Theta) \times N_D(\Theta)$ block elements. From the Appendix, the determinant of S_Θ is given by

$$|S_\Theta| = |S|^{N_D(\Theta)-1} \times |S + N_D(\Theta)\Sigma|$$

From Eq. (20) it is clear that $F(\Theta)$ is always greater than or equal to zero, so that the posterior probability of Θ increases as $F(\Theta)$ approaches zero. The first summation in Eq. (20) is a measure of how well the measured displacements of the targets under Θ match the expected bulk motion $\boldsymbol{b}$. A good match is indicated by low values for the $\boldsymbol{\nu}_i(\Theta)$, which make up the summation. The second summation is over quadratic forms in $[\boldsymbol{\nu}_k(\Theta) - \boldsymbol{\nu}_i(\Theta)]$. The vector $[\boldsymbol{\nu}_k(\Theta) - \boldsymbol{\nu}_i(\Theta)]$ is the difference between the measured displacement of target i and that of target k under hypothesis Θ. Thus, the second summation is a measure of the similarity of the displacement of each target (the summation being over all target pairs) and is independent of the expected bulk motion $\boldsymbol{b}$. Hence, the hypothesis Θ minimizing the summation $F(\Theta)$ is the result of a tradeoff between matching the prior model parameters and matching the model structure. The balance between the two summations is controlled by the relative uncertainties in the bulk motion, the within group motion and sensor errors.

A flow diagram of the algorithm is given in Fig. 2.

All models used in this analytic solution were linear and Gaussian. This limits the group bulk motion to a simple translation. Extension to allow nonlinear bulk effects such as rotation and expansion would be valuable. Sample-based techniques, such as the bootstrap approach,[8,9] offer potential approximate implementation strategies for these analytically intractable, but desirable, modeling assumptions.

VI. Hypothesis Construction and Reduction

To implement this algorithm, the measurement-target association hypotheses Θ must be enumerated. Algorithms for generating every feasible hypothesis are readily available.[10] The total number of hypotheses N_Θ can be enormous, and with substantial calculations required for each hypothesis, a full analysis can be computationally prohibitive. However, for many problems, the vast majority of these hypotheses have only very small posterior probability and, hence, make a negligible contribution to the full solution. These hypotheses, therefore, can be ignored with minimal effect (although strictly the solution would be suboptimal). It would clearly be advantageous to discard these most improbable hypotheses without evaluating their posterior probabilities. This could be viewed as a

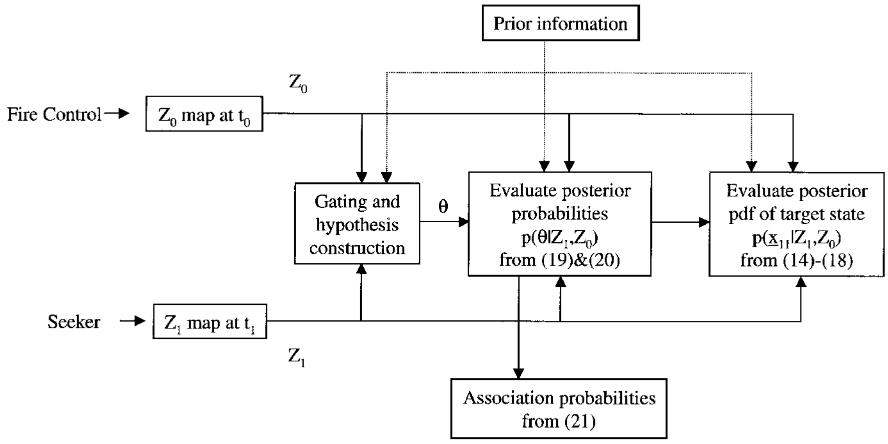

Fig. 2 Flow diagram for Bayesian algorithm (linear-Gaussian model).

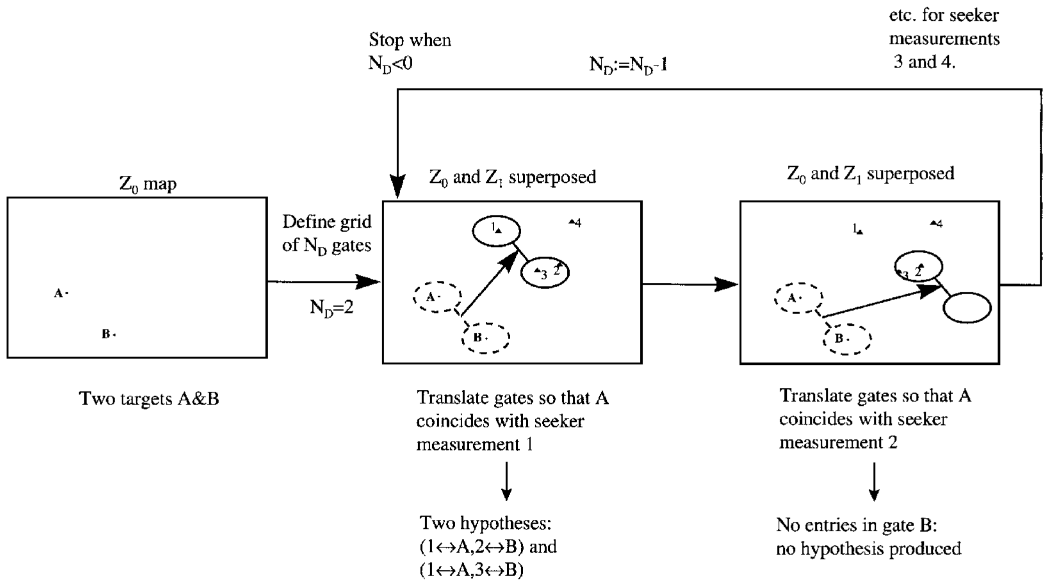

Fig. 3 Example of hypothesis construction.

coarse prefilter to select the most important hypotheses for further consideration. Such a scheme is described next. The discussion is limited to the linear-Gaussian case, with only state independent spurious measurements. Also, the prior probabilities for the hypotheses have been assumed to be uniform, i.e., uninformative, and hence it follows that hypotheses having low likelihood will also have low posterior probability.

The technique relies on the clear and complete picture of the targets obtained at time t_0. This picture effectively defines a grid of gates in the measurement space of the second sensor at time t_1. The relative spacing of the grid centers is determined by the relative spacing of the target state estimates $\hat{x}_{0i}$ (obtained from measurements at t_0). The size of the gates is chosen to accommodate (with high probability) independent and within group motion, together with the measurement errors of the two sensors. Thus, it should be possible to position the grid of gates within the measurement space, so that each measurement z_{1j} falls within one of the gates. Moreover, the translation of the grid should be compatible with the uncertainty in the bulk motion of the targets, defined by the mean $\boldsymbol{b}$ and covariance Ω. This possible translations of the grid that meet these conditions identify hypotheses Θ, which have a nonvanishing posterior probability. The reasoning leads to a fairly economical scheme for constructing a restricted set of feasible hypotheses. The proposed algorithm consists of the following steps (also see the illustrative example in Fig. 3). Repeat steps 1–5 with $N_D = 1, \ldots, \min(N_1, N)$.

1) Determine all possible sets of size N_D from N. For each of these, repeat steps 2–5 with $j = 1, \ldots, N_1$.

2) Consider the association of gate 1 with z_{1j}. If this is incompatible with the known degree of uncertainty in the bulk motion, then increment j and repeat, otherwise go to step 3.

3) Estimate the bulk motion for this association and translate the grid accordingly.

4) Test each measurement z_{1k} for membership of each gate. Thus, construct a matrix indicating which measurements fall within each of the gates.

5) Using the rule that every gate must be associated with exactly one measurement and each measurement may only be associated with (at most) one gate and that every unassociated measurement is a clutter return, construct all possible hypotheses Θ.

In this algorithm, there are N_D gates in the grid, each gate corresponding to a detected target at time t_1. Thus, there should be no empty gates. Because it is not known which of the N targets were detected at t_1, hypotheses must be constructed for every possible combination of target detections. Step 1 is a combinatorial algorithm that generates the $N!/N_D!\,(N-N_D)!$ possible grids. The algorithm NEXKSB (taken from Ref. 10, pages 21–34), which generates the combinations in alphabetical order, has been employed. The gates are numbered from 1 to N_D, and each combination is effectively a mapping from gate number to target subscript. Because the number of targets detected, N_D, is unknown, the process is repeated for all possible values of N_D. Note that the hypothesis where $N_D = 0$ is always accepted.

Step 2 considers a possible association between one of the measurements z_{1j} and the target corresponding to gate 1, which will be denoted target i. The current value of i is determined by step 1. Possible association between z_{1j} and target i is tested using a χ^2 test with threshold T_1. It is possible that the preceding association test with target i may be failed for all measurements. In this case, every combination of grid gates for which target i corresponds to gate 1 will also fail, and so need not be tested. This is easily implemented as step 1 generates combinations in alphabetical order.

Given that the association test is passed, step 3 produces an estimate of the bulk motion based on $\hat{x}_{0i}$ and z_{1j}, and the grid of gates is translated accordingly. In step 4, every measurement is tested for membership of each of the gates, again using a χ^2 test with threshold T_2. The result is stored in a matrix that indicates which of the measurements fell within each of the gates.

In the final step 5, a hypothesis construction algorithm is run over the gate membership matrix. This generates all feasible hypotheses Θ from the requirement that for every hypothesis, every gate must be associated with exactly one measurement and each measurement may only be associated with one gate. When all gates have been assigned a measurement for a given hypothesis, all remaining measurements are taken to be clutter.

The following points should be noted.

1) The hypotheses selected by this algorithm may depend on the labeling of the gates, and in particular on which target corresponds to gate 1. This is because gate 1 is used to define the translation of the grid in step 3 of the algorithm. Labeling a different gate as number 1 would result in a different set of gate translations and so might change the selection of the hypotheses. However, provided the gate sizes are chosen to be sufficiently large, the only difference should be in those hypotheses that are only barely acceptable, i.e., highly improbable, and so are of little consequence in the final solution.

2) The setting of the thresholds T_1 and T_2 should strictly depend on N_D, the number of targets detected at time t_1. Suppose that thresholds were chosen so that the probability of incorrectly rejecting a measurement was P_0. If each individual measurement-gate acceptance test were independent, then the probability of rejecting the correct hypothesis would be $P_c = 1 - (1 - P_0)^{N_D}$. Because of the coupling from the bulk motion, these tests are not independent, and so P_c is somewhat less than this upper bound. However, for fixed P_0, the value of P_c will rise with N_D. It is recommended that the thresholds T_1 and T_2 should be chosen with rejection probabilities of 0.001 (so that for two-dimensional measurement vectors $T_1 = T_2 \approx 13.8$); this should be reasonable for small N_D, for example, less than 10.

VII. Simulation Example

To illustrate the particular solution obtained in Sec. V, we present the following simulation example. The solution is obtained by evaluating Eqs. (15–18), (19), and (20). For the results given here, the posterior densities and probabilities have only been evaluated for those hypotheses selected by the construction algorithm of Sec. VI.

For this simulation example, the following assumptions are made.

1) The state vector $\boldsymbol{x}$ and the measurement vector $\boldsymbol{z}$ are two dimensional and are composed of the Cartesian coordinates of the target position in a plane.

2) $M_0 = m_0 I_2$, $R_0 = r_0 I_2$, $R_1 = r_1 I_2$, $Q = q I_2$, and $\Omega = \sigma^2 I_2$, where m_0, r_0, r_1, q, and σ^2 are positive scalars and I_2 is the 2×2 identity matrix.

3) Here $m_0 \gg r_0$.

4) $A = \Gamma_1 = \Gamma_0 = H_1 = H_0 = I_2$.

5) Exactly matched models have been assumed.

Thus, the two Cartesian coordinates of the targets are independent, and the uncertainty on each coordinate is the same. The number of targets in the initial group is six, and the targets are arranged in a triangular pattern. Figure 4 shows the target and measurement positions seen by a fire control system at targeting time t_0, with the measurements having been numbered. The actual positions of the targets are represented by the triangles and the measurements by the crosses. The prior mean for the group motion is $\boldsymbol{b} = (0.0, 0.0)^T$ with variance $\sigma^2 = 10.0$. This value of σ^2 includes uncertainty about the bulk motion of the targets, together with uncertainty in sensor grid alignment due to missile navigation errors and misalignments. The prior mean for the bulk motion has been set to zero to indicate no prior knowledge as to which direction the group will move. For

Table 1 $Pr\{\psi(i) = j \mid Z_1, Z_0\}$**: posterior probability that measurement** z_{1j} **originates from target** i

Measurement number j at t_1	Target number i 1	2	3	4	5	6
1	0.9395	0.0605	0.0000	0.0000	0.0000	0.0000
2	0.0000	0.8267	0.0000	0.1733	0.0000	0.0000
3	0.0258	0.0035	0.9677	0.0000	0.0000	0.0030
4	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000
5	0.0004	0.0000	0.0315	0.0000	0.0000	0.9680
Not detected	0.0343	0.1093	0.0008	0.8267	0.0000	0.0290

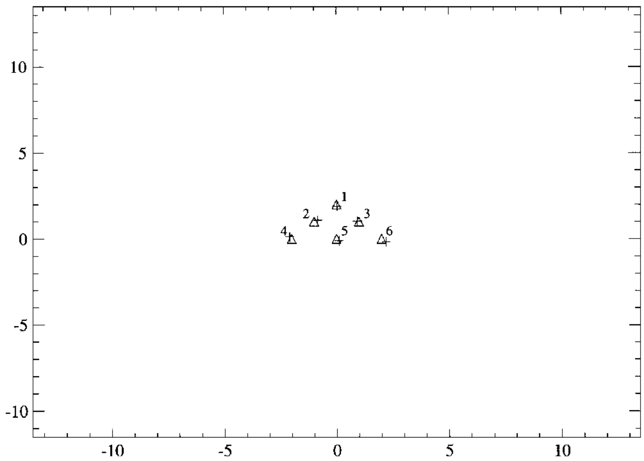

Fig. 4 Target and measurement positions at t_0**.**

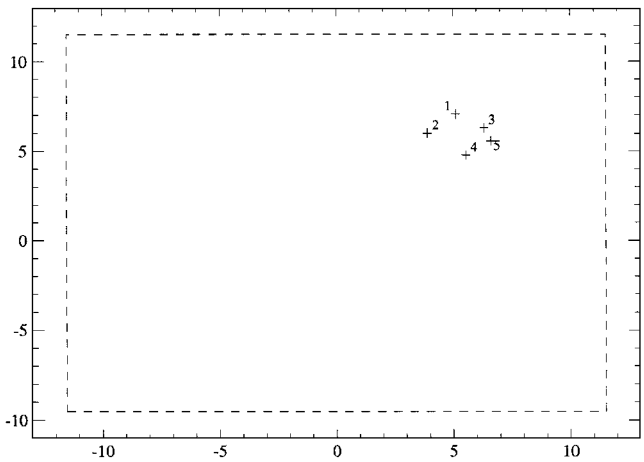

Fig. 5 Measurement positions at t_1**.**

the within group maneuvers, the variance $q = 0.2$. Measurement accuracy of the sensors is taken to be $r_0 = r_1 = 0.01$. At the time t_1 of seeker activation, the detection probability is taken to be $P_D = 0.8$. The observation region of the seeker at t_1 is set via the predictive distribution $p(X_1 \mid Z_0)$. For target i, the predictive distribution is $p(\boldsymbol{x}_{1i} \mid z_{0i}) = N[z_{0i} + \boldsymbol{b}, (q + r_0 + r_1 + \sigma^2) I_2]$. By evaluating $z_{0i} + \boldsymbol{b} \pm 3(q + r_0 + r_1 + \sigma^2)^{1/2} (1, 1)^T$ for each of the N targets in turn, and taking the maximum and minimum in the x and y directions, we obtain a box that the group falls within with high probability. The aim here is to derive an observation region sufficiently large to contain the whole target formation and not just the designated target. The resulting search box volume is $V = 497$. Two cases of this example have been simulated: with and without clutter.

First consider the no-clutter case for which the measurements at t_1 are shown in Fig. 5. In this realization, the actual bulk translation of the group was $\boldsymbol{B} = (5, 5)^T$, and the sensor observation region is indicated by the dashed rectangle. Note that only five of the six targets are detected, and it appears that target number 4 has been missed. This is the case, and the correct association Θ is indicated by the

Table 2 Top 10 components of the posterior PDF of target 1

k	$\Theta_k = (\lambda, \tau)_k$	$p(\Theta_k \mid Z_1, Z_0)$	$\hat{x}_{11}(\Theta_k)$		$P_1(\Theta_k)$		$\hat{\boldsymbol{B}}(\Theta_k)$	
1	$(1, 2, 3, 0, 4, 12)$	0.050	5.130	7.030	0.010	0.010	5.337	5.055
2	$(1, 2, 3, 0, 4, 5)$	0.050	5.114	7.035	0.010	0.010	4.990	5.169
3	$(1, 2, 3, 35, 4, 5)$	0.049	5.118	7.044	0.010	0.010	5.060	5.374
4	$(1, 2, 3, 35, 4, 12)$	0.039	5.131	7.023	0.010	0.010	5.349	5.279
5	$(1, 2, 5, 0, 4, 12)$	0.030	5.133	7.023	0.010	0.010	5.397	4.912
6	$(0, 1, 3, 2, 5, 12)$	0.030	5.995	7.454	0.254	0.254	5.969	5.568
7	$(3, 1, 5, 2, 4, 12)$	0.030	6.300	6.322	0.010	0.010	5.895	5.121
8	$(1, 0, 3, 2, 4, 12)$	0.029	5.141	7.038	0.010	0.010	5.585	5.248
9	$(0, 1, 3, 2, 4, 12)$	0.023	5.783	7.294	0.254	0.254	5.756	5.408
10	$(1, 35, 3, 2, 4, 5)$	0.017	5.118	7.044	0.010	0.010	5.060	5.374

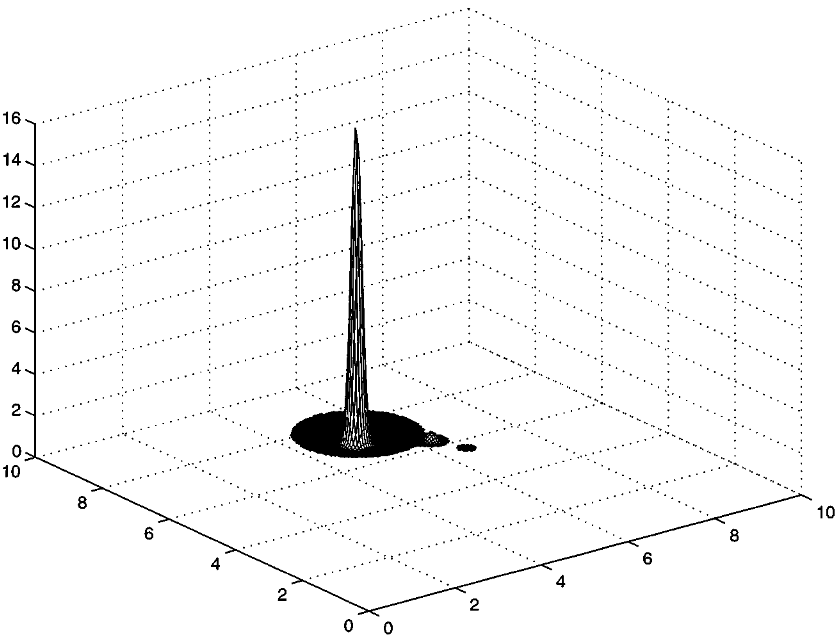

Fig. 6 Surface plot of $p(x_{11} \mid Z_1, Z_0)$.

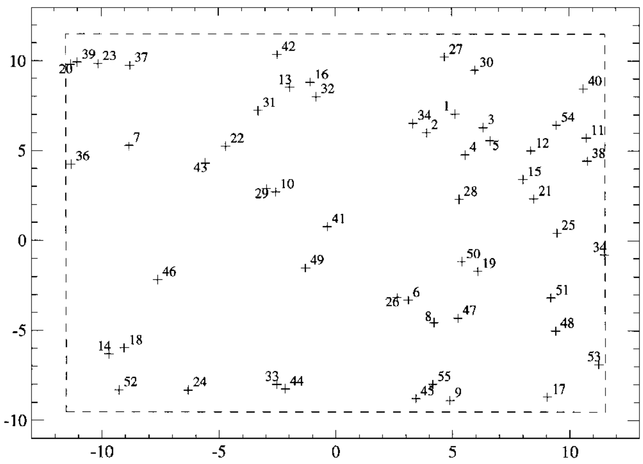

Fig. 7 Measurement positions at t_1.

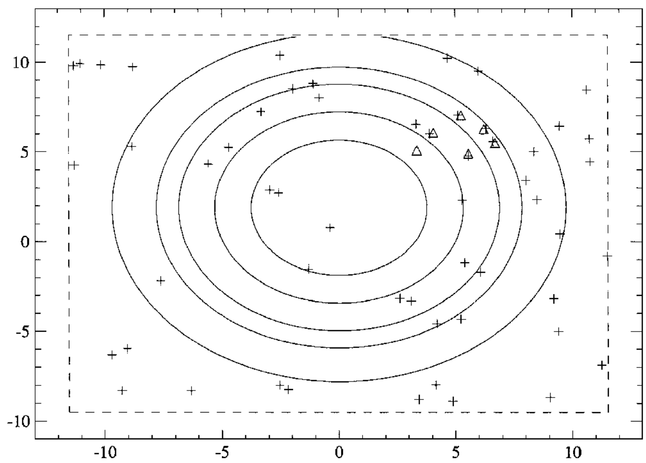

Fig. 8 Contour plot of $p(x_{11} \mid Z_0)$.

6-tuple $\Theta = (1, 2, 3, 0, 4, 5)$, [where the ith member of the 6-tuple is $\psi(i)$, the number of the measurement from Z_1 associated with target i and 0 indicates a missed target]. With the prior information that no clutter is present, the hypothesis construction algorithm generated 43 out of the $6!/(6-5)! = 720$ possible hypotheses. From these hypotheses, using the results of the preceding section, the posterior density $p(\boldsymbol{x}_{11} \mid Z_1, Z_0)$ of the designated target (number 1) may be evaluated. A surface plot of this PDF is shown in Fig. 6. As would be expected, for this fairly easy problem, the PDF exhibits a large spike over the position of the required target. The low humps around the main spike result from hypotheses that do not associate measurement 1 at t_1 with target 1. It should be noted that other useful information can be extracted from this Bayesian solution to the selection problem. In particular, the possibility that measurement j from Z_1 originates from target i may be obtained by summing the posterior probabilities of the appropriate association hypotheses:

$$Pr\{\psi(i) = j \mid Z_1, Z_0\} = \sum_{\Theta s.t. \psi(i) = j} Pr\{\Theta \mid Z_1, Z_0\} \qquad (21)$$

These probabilities have been evaluated for this example and are listed in Table 1. Note that there is a high probability that target 1 originates from z_{11} and that target 4 is not detected at t_1.

Now consider the more difficult problem, when clutter is present. Figure 7 shows measurements at time t_1 for a clutter density of $\mu = 0.1$. Thus, the expected number of clutter measurements is $\mu V = 49.7$. In fact, the simulation produced 50 clutter measurements within the observation region together with the five target detections (in the same positions as for the no-clutter case and labeled 1–5). Because it is assumed that target and clutter measurements are indistinguishable, there is clearly considerable uncertainty as to the location of the targets.

Figure 8 shows a contour plot of $p(\boldsymbol{x}_{11} \mid Z_0)$, the predictive PDF of $\boldsymbol{x}_{11}$ prior to receiving Z_1. Contours containing 50, 75, 90, 95, and 99% of the probability mass have been shown. The measurement set Z_1 has been overlaid on Fig. 8, and the target positions at t_1 are given (by triangles). Note that due to the large number of measurements in the region of high-probability density, techniques such as the PDAF[3] are inappropriate. Also note that measurement 41 is closest to the predictive mode for target 1. The hypothesis generation algorithm, with χ^2 thresholds of 13.816 ($p = 0.001$), gave 15,195 hypotheses from a total of about 2×10^{10} possible associations, a reduction of around 10^6:1. Clearly the full total is computationally prohibitive, whereas the reduced total is acceptable.

Table 2 gives some parameters of the posterior PDF. In Table 2, k is the hypothesis number, Θ_k is the hypothesis, $p(\Theta_k \mid Z_1, Z_0)$ is the posterior probability of hypothesis Θ_k, and $\hat{\boldsymbol{x}}_{11}(\Theta_k)$ and $P_1(\Theta_k)$ are the posterior mean and covariance of $p(\boldsymbol{x}_{11} \mid Z_1, Z_0, \Theta_k)$. Note that because the covariance matrix is diagonal, only the variances in the x and y directions have been given. For each hypothesis Θ_k, the posterior expected value of the bulk motion $B(\Theta_k)$ is also given. The components have been sorted in descending order of posterior probability, and so Θ_1 has the largest posterior probability. The 10 components with largest posterior probability are shown in Table 2. Although the true hypothesis is at position 2, the most probable component is in error only for target 6 (which is incorrectly hypothesised to be associated with measurement 12). The overall mean and covariance of the required mixture $p(\boldsymbol{x}_{11} \mid Z_1, Z_0)$ are $\boldsymbol{x}_{11}(\Theta) = (4.901, 6.500)^T$ and

$$P_1(\Theta) = \begin{pmatrix} 4.334 & 0.550 \\ 0.550 & 4.796 \end{pmatrix}$$

The true value of the target state is $\boldsymbol{x}_{11} = (5.250, 7.012)^T$ and of the bulk motion $\boldsymbol{B} = (5.000, 5.000)^T$.

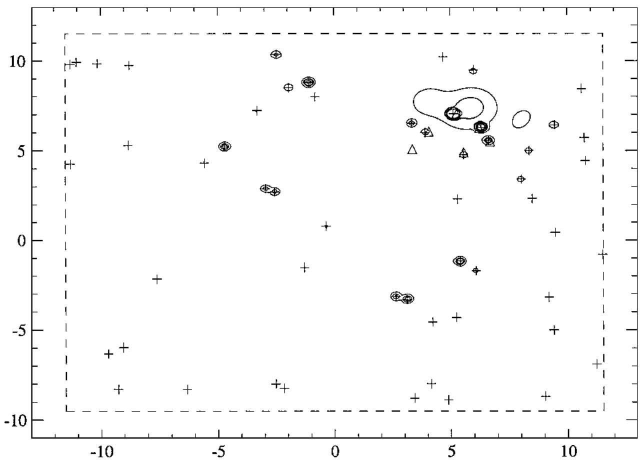

Fig. 9 Contour plot of $p(x_{11} \mid Z_1, Z_0)$.

Figure 9 shows a contour plot and Fig. 10 gives a surface plot of the posterior PDF $p(\boldsymbol{x}_{11} \mid Z_1, Z_0)$. The contour lines in Fig. 9 are set at the 50, 75, 90, 95, and 99% heights of the dominant mixture component. The many widely spaced modes indicate the difficulty of the association task; however, most probability mass is in the vicinity of the target. Although measurement 41 is closest to the predictive mode in Fig. 8, the group pattern information contained in Z_1 was sufficiently strong to move the majority of posterior probability mass away to other potential groupings. However, the key point is that this posterior PDF is a rigorous probabilistic quantification of uncertainty in the designated target location and provides the necessary information for an acquisition decision.

To provide an indication of the required computation time for this Bayesian algorithm, the execution times for the preceding two examples have been measured for a DEC Alpha EV5.6 operating at 450 MHz using 64-bit arithmetic. The first example has no clutter, and only 43 hypotheses need to be evaluated (selected from a possible 720 by the gating scheme). Execution time, including gating, was about 0.018 s. The second example is much more challenging and requires over 15,000 hypothesis evaluations (selected from about 2×10^{10} possibilities). Total execution time for this case was 0.58 s, comprising 0.28 s for gating and hypothesis construction and 0.3 s for executing the Bayesian algorithm. It should be noted that no deliberate attempt has been made to optimize the code, and it is possible that significant reduction in this computation time could be achieved. However, for real-time applications, it would be necessary to impose an upper limit on the number hypotheses via an adaptive version of the gating algorithm.

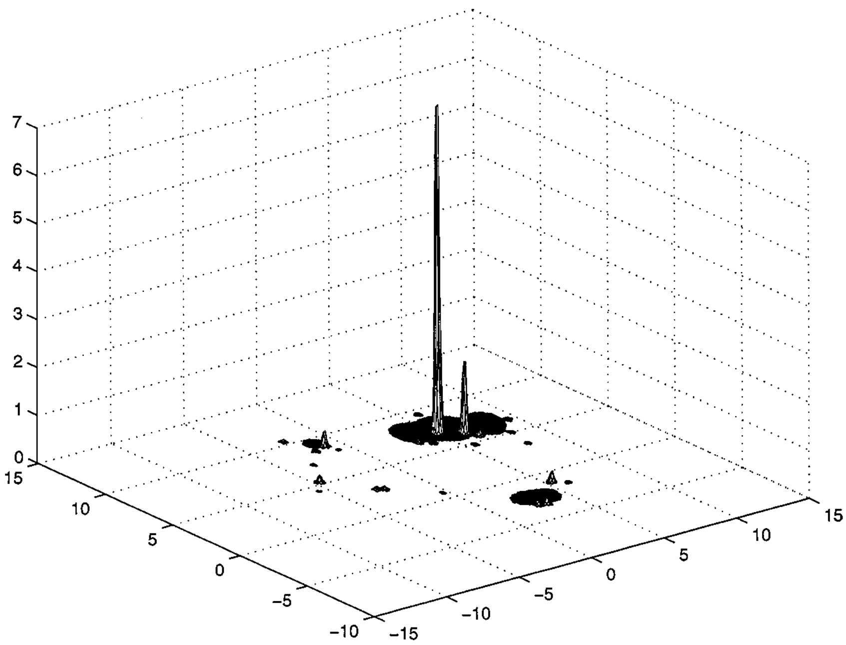

Fig. 10 Surface plot of $p(x_{11} \mid Z_1, Z_0)$.

VIII. Conclusions

The posterior PDF of the designated target's state required for acquisition is obtained as a mixture PDF resulting from the superposition of all possible measurement association hypotheses. An important feature of this Bayesian solution is that the posterior PDF precisely quantifies the uncertainty in the state of the designated target, thus allowing an informed acquisition decision. An easily identified target will result in the probability mass being concentrated about a single point in state space, whereas significant uncertainty as to the identity of the required target will result in a multimodal distribution with a number of distinct concentrations of probability mass. This can be seen in one of the simulation examples, where the presence of very dense clutter has lead to many modes in the posterior PDF.

Appendix: Proof of Matrix Determinant for Measurement Likelihood Evaluation

We first state the following result for partitioned matrices, see Ref. 11, page 258. If A and D are square matrices and A is nonsingular, then

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |A||D - CA^{-1}B|$$

Theorem: For the $n \times n$ element partitioned matrix

$$S_n(A, B) = \begin{pmatrix} A+B & B & \cdots & B \\ B & A+B & & \vdots \\ \vdots & & \ddots & B \\ B & \cdots & B & A+B \end{pmatrix}$$

the determinant is given by

$$|S_n(A, B)| = |A|^{n-1}|A + nB|$$

Proof: For $n = 1$, this is trivially true. Now assume the theorem holds for n. Then,

$$|S_{n+1}(A, B)| = \left|\begin{array}{c:ccc} A+B & B & \cdots & B \\ \hdashline B & & & \\ \vdots & & S_n(A, B) & \\ B & & & \end{array}\right|$$

$$= |A + B|\left|S_n(A, B - B(A + B)^{-1}B)\right|$$

$$= |A + B|\left|S_n(A, A(A + B)^{-1}B)\right|$$

Now, from the theorem,

$$\left|S_n(A, A(A + B)^{-1}B)\right| = |A|^{n-1}|A + nA(A + B)^{-1}B|$$

$$= |A|^{n-1}|A(A + B)^{-1}(A + (n + 1)B)|$$

$$= |A|^{n}|A + B|^{-1}|A + (n + 1)B|$$

which gives

$$|S_{n+1}(A, B)| = |A|^{n}|A + (n + 1)B|$$

The theorem is proved by induction.

Acknowledgments

The authors acknowledge many useful discussions with Adrian Smith of Imperial College, University of London. We also gratefully acknowledge useful comments from the anonymous reviewers that improved the presentation and content of the article. This research was sponsored by the UK MOD Corporate Research program TG3.

References

[1]Blackman, S. S., *Multiple Target Tracking with Radar Applications*, Artech House, Norwood, MA, 1986.

[2]Gordon, N. J., Salmond, D. J., and Everett, N. O., "Target Acquisition/Selection: A Comparison Between a Bayesian Technique and the Munkres Algorithm," *Acquisition, Tracking and Pointing XII*, edited by M. K. Masten and L. A. Stockum, Vol. 3365, Society of Photo-Optical Instrumentation Engineers, Orlando, FL, 1998, pp. 264–276.

[3]Bar-Shalom, Y., and Fortmann, T. E., *Tracking and Data Association*, Vol. 179, Academic, Orlando, FL, 1988, pp. 163–183.

[4]Pao, L. Y., "Multisensor Multitarget Mixture Reduction Algorithms for Target Tracking," *Journal of Guidance, Control, and Dynamics*, Vol. 17, No. 6, 1994, pp. 1205–1211.

[5]Mori, S., Chang, C.-Y., Tse, E., and Wishner, R. P., "Tracking and Classifying Multiple Targets Without A-Priori Identification," *IEEE Transactions on Automatic Control*, Vol. AC-31, No. 5, 1986, pp. 401–409.

[6]Kenley, C. R., and Casaletto, T. R., "Multi-Target Tracking Using Influence Diagram Models (with Discussion)," *Influence Diagrams, Belief Nets and Decision Analysis*, edited by R. M. Oliver and J. Q. Smith, Wiley, New York, 1990, Chap. 10, pp. 229–251.

[7]Gordon, N. J., Salmond, D. J., and Fisher, D., "Bayesian Target Tracking After Group Pattern Distortion," *Signal and Data Processing of Small Targets 1997*, edited by O. E. Drummond, Vol. 3163, Society of Photo-Optical Instrumentation Engineers, San Diego, CA, 1997, pp. 238–248.

[8]Gordon, N. J., "Bayesian Methods for Tracking," Ph.D. Thesis, Dept. of Statistics, Imperial College, Univ. of London, London, England, UK, Sept. 1993.

[9]Gordon, N. J., "A Hybrid Bootstrap Filter for Tracking in Clutter," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 33, No. 1, 1997, pp. 353–358.

[10]Nijenhuis, A., and Wilf, H. S., *Combinatorial Algorithms*, Academic, New York, 1975, pp. 21–34.

[11]Searle, S. R., *Matrix Algebra Useful for Statistics*, Wiley, New York, 1982, p. 258.