Set-Valued Multiple-Target Tracking

Wynn Stirling and Jeffrey Thompson
Electrical & Computer Engineering Department
CB459 Brigham Young University
Provo, UT 84602

Abstract— **The performance of a MTT (multiple target tracking) method should degrade gracefully as the conditions of the collection become less favorable to optimal operation. By stressing the avoidance, rather than the explicit minimization, of error, we obtain a decision rule for trajectory-data association that does not require the resolution of all conflicting hypotheses when the database does not contain sufficient information to do so reliably.**

I. Introduction

Multiple Target Tracking (MTT) is a time-varying joint decision and estimation problem consisting of a rule to associate sensor outputs and target vehicle trajectories and an estimator to calculate the trajectories based upon the associated sensor outputs. The solution to this problem requires the extraction from the data of as much information as possible about the number of targets and the trajectory of each. It is desirable to distinguish reliably all targets of interest from background noise, and to associate accurately each target with the available data.

We develop a trajectory-data association decision rule that is expressed in terms of two probabilities: one governing the informational value of the association hypothesis, and one governing the subjective belief, or credal, probability of the association hypothesis [2]. We also define a criterion of *serious possibility*, and all trajectory-data associations that are seriously possible are retained; it is not necessary to resolve all conflicting hypotheses before processing more data.

We employ an *set*-valued estimator, rather than the conventional point-valued estimator, to describe the evolution of the vehicle trajectory. The set-valued estimator is based upon the set-valued Kalman filter [3], which computes a convex set of trajectories, all with equal claims to validity, given the observations. If a trajectory is completely observable, the radius of the convex set decreases to zero as the quantity of data increases, and the set-valued estimate asymptotically becomes point-valued. Under less favorable circumstances, the radius of the set remains finite and may even grow, thus providing a more comprehensive characterization of the trajectory than does a single point estimate.

Manuscript received 29 July, 1991. This work was supported by ESL, Inc., Sunnyvale, CA.

II. Epistemic Utility Theory

For an inquiry under investigation, suppose there are finitely many hypotheses that may be considered. Let U denote this set of possible answers and assume that exactly one element of U is correct, and that all elements of U are consistent with our present state of knowledge. U is said to be an *ultimate partition*, and a *potential answer* is the collection of hypotheses remaining after we have rejected all members of a subset of U. Each element of a potential answer is said to be a *serious possibility*. A potential answer is degenerate if no elements of U are rejected, and we may not reject all members of U.

Epistemic utility is a probability that is composed of a convex combination of two probabilities, one measuring the importance of acquiring new information, the other measuring the importance of avoiding error. For any $g \subset U$, we define the utility of accepting g in the interest of avoiding error as $\mathcal{T}(g;\ell) = 1$ if $\ell = true$, and $\mathcal{T}(g;\ell) = 0$ if $\ell = false$. In addition to the cost of error, we also apportion a unit of informational value to each hypothesis $h_i \in U$ by assigning to elements of U non-negative real values such that their sum is unity. If we enumerate $U = \{h_1, h_2, \dots, h_n\}$, and let $\mathcal{M}(h_j) \geq 0$ denote the value assigned to h_j, then $\sum_{j=1}^{n} \mathcal{M}(h_j) = 1$, and for any set $g \subset U$, we define

$$\mathcal{M}(g) = \sum_{h_j \in g} \mathcal{M}(h_j) \tag{1}$$

as the informational value of rejecting g. The utility of accepting g in the interest of acquiring new information regardless of its truth-value is, then, $\mathcal{C}(g) = 1 - \mathcal{M}(g)$.

ISSN# 0-7803-0233-8/91 $1.00©1991 IEEE

We may address the conflict that exists between the goals of avoiding error and acquiring information by defining an *epistemic utility function* for acquiring error-free knowledge (that is, making a decision) as the convex combination $u(g;\ell) = \alpha\mathcal{T}(g;\ell) + (1-\alpha)\mathcal{C}(g)$. The quantity α represents the relative importance attached to avoiding error versus acquiring new information. We must restrict $\frac{1}{2} \leq \alpha \leq 1$ to ensure that no erroneous answer is preferred to any correct answer. Since all utility functions that are related by a positive linear transformation are equivalent, we may simplify this utility function by defining $u^a(g;\ell) = \frac{1}{\alpha}u(g;\ell) - \frac{1-\alpha}{\alpha}$. The resulting utility function for accepting g in the interest of both avoiding error and acquiring new knowledge is

$$u^a(g;\ell) = \begin{cases} 1 - b\mathcal{M}(g) & \text{if } \ell = t \\ -b\mathcal{M}(g) & \text{if } \ell = f \end{cases}, \tag{2}$$

where $b = \frac{1-\alpha}{\alpha}$ is the *coefficient of boldness.* This coefficient is constrained to lie in the interval $[0\ 1]$. The closer b is to unity, the less caution is exercised that error will be introduced (increased boldness in accepting hypotheses); the closer b is to zero, the lower the risk of error (decreased boldness).

We must also establish a probabilistic measure of belief for the hypotheses that are available. *Credal* probability is probability formed on the basis of subjective judgment, represents the likelihood that an option is true, and is independent of any informational value or demand that might be associated with the option.

For a given ultimate partition U, let $\mathcal{Q}(g)$ denote the credal probability assignment to any element $g \subset U$. For $g \subset U$, the expected utility is

$$\begin{aligned} E_Q u_a(g;\ell) &= [1 - b\mathcal{M}(g)]\mathcal{Q}(g) - b\mathcal{M}(g)[1 - \mathcal{Q}(g)] \\ &= \mathcal{Q}(g) - b\mathcal{M}(g), \end{aligned}$$

where $E_{\mathcal{Q}}(\cdot)$ is mathematical expectation. This expected utility represents a tradeoff between the desire to acquire new knowledge and the desire to avoid error. The choice of b establishes a threshold at which the demand for knowledge renders the risk of error worthwhile.

We may adopt any set of hypotheses in the Boolean algebra generated by the elements of the ultimate partition, U. This expands our possibilities; we are not constrained to select only the elementary hypotheses, h_i, but may choose any subset of them. This decision philosophy may be summarized as follows:

Levi's Rule of Expected Utility [1, page 53]: *Given a finite ultimate partition U, an information-determining probability function $\mathcal{M}$ defined over the Boolean algebra of elements of U, an expectation-determining probability function $\mathcal{Q}$ defined over the same algebra, and an index of boldness b, the agent should reject all and only those elements of $h_i \in U$ satisfying $\mathcal{Q}(h_i) < b\mathcal{M}(h_i)$.*

III. Multiple Target Tracking

A. Track Initialization

We assume that each track is characterized by a linear stochastic dynamics model of the form

$$\mathbf{x}_{jt} = F_{jt-1}\mathbf{x}_{jt-1} + G_{jt-1}\mathbf{u}_{jt-1}, \tag{3}$$

where $j = 1, 2, \cdots, \mathcal{T}_t$ and $\mathbf{u}_{jt-1} \sim \mathcal{N}(0, Q_{jt-1})$ (i.e., $\mathbf{u}_{it-1}$ is Gaussian with mean 0 and covariance matrix Q_{jt-1}); there are $\mathcal{T}_t$ active tracks at time t ($\mathcal{T}_t$ is unknown). The number of tracks is allowed to vary since tracks may be initiated or terminated at any time. We assume that all tracks lie in the same state space.

In the interest of brevity, we shall restrict attention to the outputs of a single sensor, and assume that this output may be characterized by a linear stochastic model of the form

$$\mathbf{z}_{it} = H_{it}\mathbf{x}_{jt} + \mathbf{v}_{it}, \quad i = 1, \cdots s_t, \tag{4}$$

where s_t is the number of observations vectors at time t and $\mathbf{z}_{it}$ is an r_{it}-dimensional random vector, and $\mathbf{v}_{it} \sim \mathcal{N}(0, R_{it})$. Each observation vector, therefore, lies in an r_{it} dimensional space corresponding to the column space of H_{it}. We assume that each such space is a subset of the state space ($r_{it} \leq n$). We do not, however, require all observations to lie in the same dimensional subspace of the state space, and we permit the dimensionality of these column spaces to be time-varying.

Before data are collected, we characterize the target environment with one set-valued track defined by an initial credal matrix K_0, an initial centroid state $\underline{c}_0$, and a prior covariance matrix Π_0. The initial state-vector set is

$$\mathbf{X}_{0|0} = \left\{\mathbf{x} \sim \mathcal{N}(\underline{x}, P_{0|0}) \,:\, \underline{x} \in \underline{X}_{0|0}\right\},$$

where

$$\underline{X}_{0|0} = \left\{\underline{x} \in \Re^n : (\underline{x} - \underline{c}_{0|0})^T \left[S_{0|0}\right]^{-1} (\underline{x} - \underline{c}_{0|0}) \leq 1\right\},$$

$P_{0|0} = \Pi_0$ is a positive-definite matrix, and $S_{0|0} = K_0 K_0^T$. For sample times $t = 1, 2, \cdots$, we observe s_t observations $\underline{Z}_t = \{\underline{z}_{1t}, \cdots, \underline{z}_{s_t t}\}$. Let us suppose, at time t, that we have τ_{t-1} sets of predicted (from time $t-1$) random-variables of the form

$$\mathbf{X}^j_{t|t-1} = \left\{\mathbf{x} \sim \mathcal{N}(\underline{x}, P^j_{t|t-1}) \;:\; \underline{x} \in \underline{X}^j_{t|t-1}\right\},$$

where

$$\underline{X}^j_{t|t-1} = \left\{\underline{x} \in \Re^n : (\underline{x} - \underline{c}^j_{t|t-1})^T \left[S^j_{t|t-1}\right]^{-1} (\underline{x} - \underline{c}^j_{t|t-1}) \leq 1\right\}, \tag{5}$$

where $S^j_{t|t-1} = \left[K^j_{t|t-1}\right]\left[K^j_{t|t-1}\right]^T$. with

$$\underline{c}^j_{t|t-1} = F_{t-1}\underline{c}^j_{t-1|t-1} \quad (6)$$

$$P^j_{t|t-1} = F_{t-1}P^j_{t-1|t-1}F^T_{t-1} + G_{t-1}Q_{t-1}G^T_{t-1} \quad (7)$$

$$K^j_{t|t-1} = F_{t-1}K^j_{t-1|t-1}, \quad (8)$$

for $j = 1, \cdots, \tau_{t-1}$. We shall use the same F_t, G_t, and Q_t matrices for each track, thus it will not be necessary to index these matrices with their track identifications.

B. The Ultimate Partition

At each time, t, there exist s_t observation sample vectors, $\{\underline{z}_{it}\}_{i=1}^{s_t}$ of the random variables $\{\mathbf{z}_{it}\}_{i=1}^{s_t}$, and τ_{t-1} predicted track sets, $\{\underline{X}^j_{t|t-1}\}_{j=1}^{\tau_{t-1}}$. We wish to make decisions regarding the association of each $\underline{z}_{it}$ with each track set $\underline{X}^j_{t|t-1}$. We desire to apply Levi's decision-making methodology to this problem and will, therefore, adopt the strategy of accepting all track/data associations that cannot be rejected on the basis of Levi's rule of epistemic utility.

We must define an ultimate partition for each sample vector, resulting in a set of s_t ultimate partitions of the form $U_{it} = \{h_{it1}, \cdots, h_{it\tau_{t-1}}, h_{it\tau_t}\}$, $i = 1, \cdots, s_t$, where

$$h_{itj} = \begin{cases} \exists \underline{x} \in \underline{X}^j_{t|t-1} \text{ such that } \underline{z}_{it} = H_{it}\underline{x} & j = 1, \cdots, \tau_{t-1} \\ \emptyset & j = \tau_t \end{cases},$$

where $\emptyset$ signifies the hypothesis that none of the track sets associate with the observation $\underline{z}_{it}$. We shall say that track set $\underline{X}^j_{t|t-1}$ is *associated* with observation sample vector $\underline{z}_{it}$ if we fail to reject the hypothesis h_{itj}. Each ultimate partition, U_{it}, has the property that exactly one element is true, although each is logically possible. According to Levi's theory of expected utility, we may reject only those members of the ultimate partition that are not seriously possible. We do not insist that the decision that one and only one element of U_{it} be chosen as the association decision.

C. Calculation of the $\mathcal{M}$-Function

The $\mathcal{M}$-function is intended to measure the information-value of rejecting an association, rather than the truth-value of an association. A measure of the information value of rejecting the association of a predicted track set $\underline{X}^j_{t|t-1}$ with an observation $\underline{z}_{it}$ is the distance between sample values of the observation and and the track set; if the distance is small there is little value in rejecting the association (in other words, there is great value in accepting it). For $\underline{x} \in \underline{X}^j_{t|t-1} \subset \Re^n$ and $\underline{z}_{it} \in \Re^{r_{it}}$, we define the *generalized distance* between them as

$$d(\underline{x}, \underline{z}_{it}) = \|H_{it}\underline{x} - \underline{z}_{it}\| \stackrel{\text{def}}{=} \left[(H_{it}\underline{x} - \underline{z})^T(H_{it} - \underline{z}_{it})\right]^{\frac{1}{2}}.$$

We desire to normalize this distance to permit its interpretation as a probability density function. This normalization is accomplished by defining a region of the column space of H_{it} which we may assume contains all predicted observation values that may be feasibly associated with the given value for $\underline{z}_{it}$, and restricting attention only to this region, termed the *seriously possible region* for $\underline{z}_{it}$, which we shall denote as $\underline{\Sigma}_{it}$, a convex set centered at $\underline{z}_{it}$ of the form

$$\underline{\Sigma}_{it} = \left\{\underline{\zeta} \in \Re^{r_{it}} : |\zeta_\ell - z_{it\ell}| \leq \theta\rho_{i\ell}, \ell = 1, r_{it}\right\}, \quad (9)$$

where $R_{it} = \text{diag}\{\rho^2_{i1}, \cdots, \rho^2_{ir_{it}}\}$. and θ a given constant.

Since we wish $\mathcal{M}$ to be a probability, we require that the distance function, $d(\underline{x}, \underline{z}_{it})$, be normalized, thereby admitting the interpretation as an information-determining probability density function. We must normalize this function by the seriously possible region; that is, for fixed $\underline{z}_{it}$, let

$$m_{it}(\underline{x}) = \frac{d(\underline{x}, \underline{z}_{it})}{\int_{\underline{\Sigma}_{it}} \|\underline{\zeta} - \underline{z}_{it}\| d\underline{\zeta}} = \frac{\|H_{it}\underline{x} - \underline{z}_{it}\|}{\int_{\underline{\Sigma}_{it}} \|\underline{\zeta} - \underline{z}_{it}\| d\underline{\zeta}} \quad (10)$$

for all $\underline{x}$ whose projection onto the space spanned by the columns of H_{it} lies in Σ_{it}. We shall denote this space by $\Xi_{it} = \{\underline{x} \in \Re^n : \mathcal{P}_{it}\underline{x} \in \Sigma_{it}\}$, where $\mathcal{P}_{it}$ is the projection operator onto the space spanned by the columns of H_{it}.

The function $m_{it}(\underline{x})$ may be viewed as a normalized distance between $\underline{z}_{it}$ and $\underline{x} \in \underline{X}^{ij}_{t|t-1}$, and is a measure of the information gained by rejecting the association of $\underline{x} \in \underline{X}^j_{t|t-1}$ and $\underline{z}_{it}$. Intuitively, the greater the distance between the predicted observation and the actual observation, the greater the value in rejecting the hypothesis that the predicted state estimate and the observation are associated. Let $\underline{B}_{\underline{x}}$ be a ball with center at $\underline{x} \in \underline{X}^j_{t|t-1}$. The information value of rejecting the association of $\underline{B}_{\underline{x}}$ with $\underline{z}_{it}$ is

$$\mathcal{M}_{it}(\underline{B}_{\underline{x}}) = \int_{\mathcal{P}_{it}\underline{B}_{\underline{x}} \cap \Xi_{it}} m_{it}(\underline{\xi}) d\mathcal{P}_{it}\underline{\xi}, \quad , i = 1, \cdots, s_t, \quad (11)$$

where $\mathcal{P}_{it}\underline{B}_{\underline{x}}$ is the projection of the ball $\underline{B}_{\underline{x}}$ onto the space spanned by the columns of H_{it}. The vector $\mathcal{P}_{it}\underline{x}$ is the projection of $\underline{x}$ onto the same column space. We emphasize that the informational-value determining probability places all of the probability mass in the column space of H_{it}. This result is appropriate, since there is no way of assessing the informational value of rejecting track/data associations by means of components of the state that lie in the subspace orthogonal to the column space of H_{it}.

For $\underline{x} \in \underline{X}^j_{t|t-1}$ and $\underline{x}' \in \underline{X}^j_{t|t-1}$ with $\underline{x} \neq \underline{x}'$, let the diameter of the balls $\underline{B}_{\underline{x}}$ and $\underline{B}_{\underline{x}'}$ become arbitrarily small. Then the condition $\mathcal{M}_{it}(\underline{B}_{\underline{x}}) < \mathcal{M}_{it}(\underline{B}_{\underline{x}'})$ indicates that the information value of accepting the association of $\underline{x}$ with $\underline{z}_{it}$ is greater than the value of accepting the association of $\underline{x}'$ and $\underline{z}_{it}$.

D. *Calculation of the Q-Functions*

The credal probability, or Q-function, is the probability that a given track-data association is correct. This probability is a function of the statistical descriptions of the target dynamics and of the observation errors, and is characterized by the conditional distribution of the observation given the track. For each $j = 1, \cdots, \tau_{t-1}$, the set of predicted random vectors is given by

$$\mathbf{X}^j_{t|t-1} = \left\{ \mathbf{x} \sim \mathcal{N}(\underline{x}, P^j_{t|t-1}) \; : \; \underline{x} \in \underline{X}^j_{t|t-1} \right\}.$$

For each observation random vector $\mathbf{z}_{it}$ we represent the corresponding set of filtered random vectors by

$$\mathbf{X}^{ij}_{t|t} = \left\{ \hat{\mathbf{x}} \; : \; \hat{\mathbf{x}} = \mathbf{x} + W_{it}[\mathbf{z}_{it} - H_{it}\mathbf{x}], \quad \mathbf{x} \in \mathbf{X}^j_{t|t-1} \right\},$$

where W_{it} is the Kalman gain matrix given via the Kalman filter.

The probability distributions of $\{\hat{\mathbf{x}} \in \mathbf{X}^{ij}_{t|t}\}$ are characterized by the family of posterior distributions $\left\{ p^{ij}_{\hat{\mathbf{x}}_{t|t}}(\underline{\xi}; \underline{x}); \underline{x} \in \underline{X}^j_{t|t-1} \right\}$ obtained via Bayes rule. We shall restrict consideration to the case where $\mathbf{x} \in \mathbf{X}^j_{t|t-1}$ and $\mathbf{z}_{it}$ are jointly normal. Under this hypothesis, the posterior density, $p^{ij}_{\hat{\mathbf{x}}_{t|t}}(\underline{\xi}; \underline{x})$ is also normal, and it is well known that the mean and variance associated with this conditional distribution is given by the Kalman filter. For each observation $\mathbf{z}_{it} = \underline{z}_{it}$, we calculate the filtered set-valued estimate according to

$$\underline{X}^{ij}_{t|t} = \left\{ \underline{x} \in \Re^n : (\underline{x} - \underline{c}^{ij}_{t|t})^T \left[S^{ij}_{t|t} \right]^{-1} (\underline{x} - \underline{c}^{ij}_{t|t}) \leq 1 \right\}, \tag{12}$$

for $i = 1, \cdots, s_t$, where $S^{ij}_{t|t} = K^{ij}_{t|t} \left[K^{ij}_{t|t} \right]^T$ and

$$\begin{aligned} \underline{c}^{ij}_{t|t} &= \underline{c}^j_{t|t-1} + W_{ijt}[\underline{z}_{it} - H_{it}\underline{c}^j_{t|t-1}] && (13) \\ P^{ij}_{t|t} &= [I - W_{ijt}H_{it}]P^j_{t|t-1} && (14) \\ K^{ij}_{t|t} &= [I - W_{ijt}H_{it}]K^j_{t|t-1} && (15) \end{aligned}$$

for $j = 1, \cdots, \tau_{t-1}$, with W_{ijt} the Kalman gain defined by $W_{ijt} = P^j_{t|t-1}H^T_{it} \left[H_{it}P^j_{t|t-1}H^T_{it} + R_{it} \right]^{-1}$. The posterior densities assume the form

$$p^{ij}_{\hat{\mathbf{x}}_{t|t}}(\underline{\xi}; \underline{x}) = \mathcal{N}\left(\underline{x} + W_{ijt}(\underline{z}_{it} - H_{it}\underline{x}), [I - W_{ijt}H_{it}]P^j_{t|t-1} \right), \tag{16}$$

for $\underline{x} \in \underline{X}^j_{t|t-1}$. But

$$\mathcal{N}\left(\underline{x} + W_{ijt}(\underline{z}_{it} - H_{it}\underline{x}), [I - W_{ijt}H_{it}]P^j_{t|t-1} \right) = \mathcal{N}\left(\hat{\underline{x}}, P^{ij}_{t|t} \right),$$

where we define

$$\hat{\underline{x}} = \underline{x} + W_{ijt}(\underline{z}_{it} - H_{it}\underline{x}). \tag{17}$$

The posterior density is

$$p^{ij}_{\hat{\mathbf{x}}_{t|t}}(\underline{\xi}; \underline{x}) = \frac{1}{(2\pi)^{\frac{n}{2}} \left| P^{ij}_{t|t} \right|^{\frac{1}{2}}} e^{-\frac{1}{2}(\underline{\xi} - \hat{\underline{x}})^T [P^{ij}_{t|t}]^{-1} (\underline{\xi} - \hat{\underline{x}})}, \tag{18}$$

with $\hat{x}$ a function of x given by (17).

The family of densities $\{p^{ij}_{\hat{\mathbf{x}}_{t|t}}(\underline{\xi}; \underline{x}), \quad \underline{x} \in \underline{X}^j_{t|t-1}\}$ provides a measure of the truth-value of the association of each $\underline{x} \in \underline{X}^j_{t|t-1}$ with the observation $\underline{z}_{it}$. We may view this family of densities as credal probability densities, and interpret them as characterizing the belief that the above association is correct.

Let $\underline{B}_{\underline{x}}$ be a ball with center at $\underline{x}$. The credal probability that the true state lies in $\underline{B}_{\underline{x}}$ conditioned on the observational value $\underline{z}_{it}$ is

$$Q_{ijt}(\underline{B}_{\underline{x}}; \underline{x}) = \int_{\underline{B}_{\underline{x}}} p^{ij}_{\hat{\mathbf{x}}_{t|t}}(\underline{\xi}; \underline{x}) d\underline{\xi}. \tag{19}$$

For $\underline{x} \in \underline{X}^j_{t|t-1}$ and $\underline{x}' \in \underline{X}^j_{t|t-1}$ with $\underline{x} \neq \underline{x}'$, the condition $Q_{ijt}(\underline{B}_{\underline{x}}; \underline{x}) > Q_{ijt}(\underline{B}_{\underline{x}'}; \underline{x}')$ indicates that the association of $\underline{x}$ with $\underline{z}_{it}$ is more credible, or believable, than the association of $\underline{x}'$ and $\underline{z}_{it}$.

E. *The Association Likelihood Ratio Test*

The association decision problem is to determine whether or not the track $\underline{X}^j_{t|t-1}$ is associated with the observation $\underline{z}_{it}$. We shall assume a conservative attitude, and say that the entire track set $\underline{X}^j_{t|t-1}$ is associated with $\underline{z}_{it}$ if any $\underline{x} \in \underline{X}^j_{t|t-1}$ is associated with $\underline{z}_{it}$.

The decision rule may be formulated in terms of the information-determining probability density, $m_{it}(\underline{x})$ (where $\underline{x} \in \underline{X}^j_{t|t-1}$), and the family of subjective belief, or credal, probability density functions functions, $\{p^{ij}_{\hat{\mathbf{x}}_{t|t}}(\underline{\xi}; \underline{x}), \underline{x} \in \underline{X}^j_{t|t-1}\}$. We desire to apply Levi's rule of expected utility to this problem.

For $\underline{x} \in \underline{X}_{t|t-1}$, let $\underline{B}_{\underline{x}}$ be a ball with center at $\underline{x}$. Using (11) and (19), Levi's rule of expected utility indicates we may not reject the association of $\underline{B}_{\underline{x}}$ and $\underline{z}_{it}$ if

$$Q_{ijt}(\underline{B}_{\underline{x}}; \underline{x}) \geq b\mathcal{M}_{it}(\underline{B}_{\underline{x}}). \tag{20}$$

Now let the radius of $\underline{B}_{\underline{x}}$ go to zero, and define the function

$$q_{ijt}(\underline{x}) \stackrel{\text{def}}{=} p^{ij}_{\hat{\mathbf{x}}_{t|t}}(\underline{x}; \underline{x}). \tag{21}$$

Since the densities are continuous at $\underline{x}$, a necessary condition for (20) to hold for all balls $\underline{B}_{\underline{x}}$ is that

$$q_{ijt}(\underline{x}) \geq bm_{it}(\underline{x}). \tag{22}$$

If (22) holds for any $\underline{x} \in \underline{X}^j_{t|t-1}$, we may not reject the association of $\underline{x}$ with $\underline{z}_{it}$; consequently, we may not reject the association of the track set $\underline{X}^j_{t|t-1}$ with $\underline{z}_{it}$.

Since

$$p_{\hat{\mathbf{x}}_{t|t}}(\underline{\xi};\underline{x}) = \mathcal{N}\{\underbrace{\underline{x} + W_{ijt}(\underline{z}_{it} - H_{it}\underline{x})}_{\hat{\underline{x}}}, \underbrace{[I - W_{ijt}H_{it}]P^j_{t|t-1}}_{P_{t|t}}\},$$

we have

$$\begin{aligned} q_{ijt}(\underline{x}) &= \left.\frac{1}{(2\pi)^{\frac{n}{2}}|P^{ij}_{t|t}|^{\frac{1}{2}}} e^{-\frac{1}{2}(\underline{\xi} - \hat{\underline{x}})^T[P^{ij}_{t|t}]^{-1}(\underline{\xi} - \hat{\underline{x}})}\right|_{\underline{\xi}=\underline{x}} \\ &= \frac{e^{-\frac{1}{2}(H_{it}\underline{x} - \underline{z}_{it})^T W^T_{ijt}[P^{ij}_{t|t}]^{-1}W_{ijt}(H_{it}\underline{x} - \underline{z}_{it})}}{(2\pi)^{\frac{n}{2}}|P^{ij}_{t|t}|^{\frac{1}{2}}} \end{aligned} \tag{23}$$

Given the information-value determining probability density (10) and the credal probability density (23), we may form the Association Likelihood Ratio Test (ALRT):

Let $\underline{X}^j_{t|t-1}$ be a predicted track set, and let $\underline{z}_{it}$ be a sample value of the observation vector $\mathbf{z}_{it}$. We shall say that $\underline{X}^j_{t|t-1}$ and $\underline{z}_{it}$ are associated with boldness b *if $q_{ijt}(\underline{x}) \geq bm_{it}(\underline{x})$ for some $\underline{x} \in \Xi_{it} \cap \underline{X}^j_{t|t-1}$. That is, there exists $\underline{x} \in \Xi_{it} \cap \underline{X}^j_{t|t-1}$ such that*

$$\frac{1}{(2\pi)^2|P^{ij}_{t|t}|^{\frac{1}{2}}} e^{-\frac{1}{2}(H_{it}\underline{x} - \underline{z}_{it})^T W^T_{ijt}[P^{ij}_{t|t}]^{-1}W_{ijt}(H_{it}\underline{x} - \underline{z}_{it})} \geq b\frac{||H_{it}\underline{x} - \underline{z}_{it}||}{\int_{\Sigma_{it}} ||H_{it}\underline{\xi} - \underline{z}_{it}||d\underline{\xi}}, \tag{24}$$

in which case, $\underline{X}^{ij}_{t|t}$ obtained via (12) is a filtered track set. If $\underline{X}^j_{t|t-1} \cap \Xi_{it} = \emptyset$, then we deem $\underline{z}_{it}$ and $\underline{X}_{t|t-1}$ to be dissociated, *and the set $\underline{X}^{ij}_{t|t}$ has no meaning and is discarded.*

If $\underline{X}^j_{t|t-1}$ survives the ALRT for $\underline{z}_{it}$, then the likelihood that $\underline{X}^j_{t|t-1}$ is associated with $\underline{z}_{it}$ is greater than the information value gained by rejecting the association. If $b = 1$, then the decision strategy is maximally bold in the sense that as many associations $(\underline{X}^j_{t|t-1}, \underline{z}_{it})$ will be rejected as possible. As b approaches zero, the decision strategy is maximally cautious, and the likelihood diminishes that a correct association (indeed, any association) will be rejected.

IV. Example

To illustrate the concept of the ALRT methodology, consider the case of tracking targets constrained to planar motion [4]. Let $\underline{x} = [x, y, \dot{x}, \dot{y}]^T$ denote the kinematic state of a target in some convenient coordinate system ($n = 4$). The dynamics equation is

$$\mathbf{x}_t = \begin{bmatrix} 1 & 0 & \Delta_t & 0 \\ 0 & 1 & 0 & \Delta_t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{x}_{t-1} + \begin{bmatrix} \mathrm{u}_{\mathrm{x}_t} \\ \mathrm{u}_{\mathrm{y}_t} \\ \mathrm{u}_{\dot{\mathrm{x}}_t} \\ \mathrm{u}_{\dot{\mathrm{y}}_t} \end{bmatrix},$$

where Δ_t is the sample interval, and we have set $G \equiv I$. Possible target maneuvers are assumed to be characterized by the process noise, $\mathbf{u}_t$, whose covariance is Q_t. We assume that angle-of-arrival data are available; further, we assume that the sensor is sufficiently far from the target that the linearized model is adequate. For convenience we also assume that the coordinate system is resolved along the azimuth and elevation angles $H_{it} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$; that is, $r_{it} \equiv 2$. Let $\underline{x} = [x_1, x_2, x_3, x_4]^T \in \underline{X}^j_{t|t-1}$, $\underline{z}_{it} = [z_{it1},\ z_{it2}]^T$, and suppose $R_{it} = \mathrm{diag}\{\rho_1^2, \rho_2^2\}$. Then the seriously possible region is

$$\Sigma_{it} = \left\{ \begin{bmatrix} \zeta_1 \\ \zeta_2 \end{bmatrix} : |\zeta_1 - z_{it1}| \leq \theta\rho_1, \quad |\zeta_2 - z_{it2}| \leq \theta\rho_2 \right\},$$

and the information-value determining probability density is

$$m_{it}(\underline{x}) = \frac{\sqrt{(x_1 - z_{it1})^2 + (x_2 - z_{it2})^2}}{\int_{-\theta\rho_2}^{\theta\rho_2}\int_{-\theta\rho_1}^{\theta\rho_1} \sqrt{\zeta_1^2 + \zeta_2^2}d\zeta_1 d\zeta_2}. \tag{25}$$

Each $\underline{x} \in \underline{X}^j_{t|t-1}$ represents the mean value of a predicted conditional distribution. For each such $\underline{x}$ there exists a filtered conditional density of the form

$$p^{ij}_{\hat{\mathbf{x}}_{t|t}}(\underline{\xi};\underline{x}) = \mathcal{N}\{\underline{x} + W_{ijt}(\underline{z}_{it} - H_{it}\underline{x}), [I - W_{it}H_{it}]P_{t|t-1}\}, \tag{26}$$

$\underline{x} \in \underline{X}^j_{t|t-1}$, and evaluating this expression at $\underline{\xi} = \underline{x}$,

$$q_{ijt}(\underline{x}) = \frac{(2\pi)^2|P^{ij}_{t|t}|^{\frac{1}{2}}}{e^{-\frac{1}{2}[x_1 - z_{it1},\ x_2 - z_{it2}]W^T_{ijt}[P^{ij}_{t|t}]^{-1}W_{ijt}\begin{bmatrix} z_{it1} - x_1 \\ z_{it2} - x_2 \end{bmatrix}}}. \tag{27}$$

The ALRT is: Associate $\underline{X}^j_{t|t-1}$ and $\underline{z}_{it} = [z_{it1}, z_{it2}]^T$ if, for any $\underline{x} = [x_1, x_2, x_3, x_4]^T \in \underline{X}_{t|t-1}$, $q_{ijt}(\underline{x}) \geq bm_{it}(\underline{x})$.

Figure 1(a) illustrates a family of three crossing trajectories. The tracks move generally from left to right at time increases; the lines correspond to the x and y position components and the + symbols correspond to noise-corrupted observations. Figure 1(b) displays the filtered track sets corresponding to this simulation, with the ellipses representing the projections of the track sets onto position space. The initial predicted track set $\mathcal{X}_{0|-1}$ includes the entire field of view, and is not shown. The three large elliptical regions correspond to the the filtered track sets after the first set of observations have been processed. As time increases, the size of these track sets decreases rapidly; for the first few observations corresponding to Tracks 1 and 3, however, there are multiple associations, since the tracks are fairly close and the track sets are still fairly large. As more confidence is obtained in the associations, these tracks become uniquely associated, and the elliptical regions decrease rapidly in size, and will asymptotically become point tracks. At sample ten, Tracks 1 and 2 nearly intersect, and both tracks associate with the observations. These multiple associations persist for a few samples, but as the tracks diverge, the associations again become unique. Track 2 is unambiguously associated and estimated, as evidenced by the radius of the track set converging to zero, and the set-valued estimates asymptotically become point-valued.

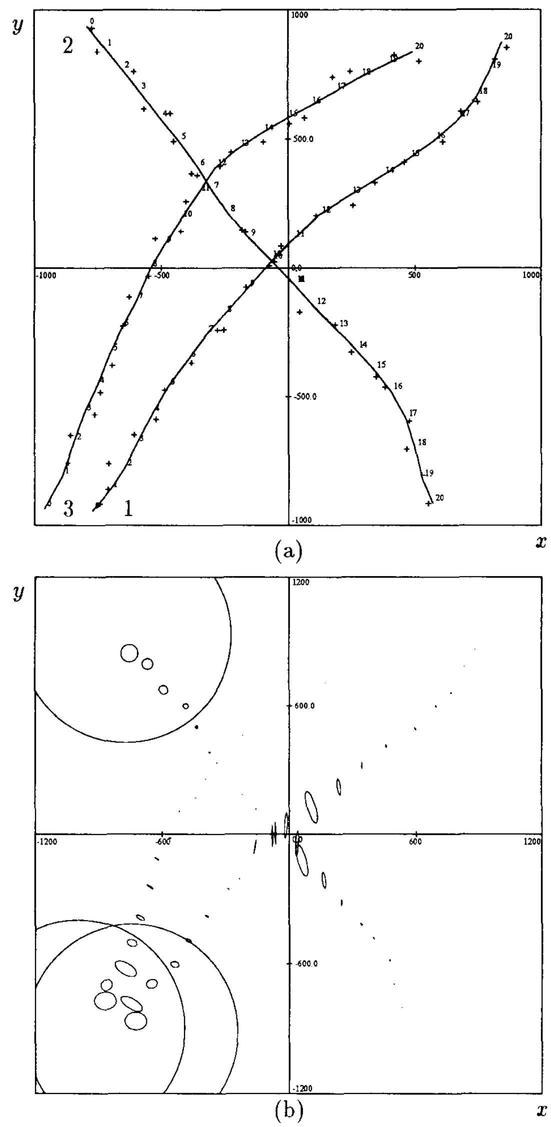

Figure 1: Three Crossing Tracks: (a) simulated trajectories and observations; (b) filtered track sets.

V. Conclusion

The use of epistemic utility, combined with set-valued estimation, makes it possible to design a multiple target tracking procedure that extracts as much information as possible from a given database without attempting to squeeze more information from the database than it contains.

References

[1] I. Levi, *The Enterprise of Knowledge*, MIT Press, Cambridge, Massachusetts, 1980.

[2] W. C. Stirling and D. R. Morrell, Convex Bayes Decision Theory, *IEEE Transactions on Systems, Man, and Cybernetics*, **21**(1):173–183, January/February 1991.

[3] D. R. Morrell and W. C. Stirling, Set-Valued Filtering and Smoothing, *IEEE Transactions on Systems, Man, and Cybernetics*, **21**(1):184–193, January/February 1991.

[4] W. C. Stirling, D. R. Morrell, G. R. Morell, and J. B. Thompson, Convex Bayes Decision Theory and Set-Valued Estimation: A New Approach to Multiple Target Tracking, Technical Report TR-S107-91.1, Electrical and Computer Engineering Department, Brigham Young University, Provo, Utah, 1991.