Discrimination Gain to Optimize Detection and Classification

Keith Kastella *

UNISYS Government Systems Group
P. O. Box 64525
St. Paul, Minnesota 55164-0525
(612) 456-2537
kastella@sp.paramax.com

February 14, 1995

Abstract

A method for managing agile sensors to optimize detection and classification based on discrimination gain is presented. Expected discrimination gain is used to determine threshold settings and search order for a collection of discrete detection cells. This is applied in a low signal-to-noise environment where target-containing cells must be sampled many times before a target can be detected or classified with high confidence. The goal of sensor management is interpreted to be to direct sensors to optimize the probability densities produced by a data fusion system that they feed. The use of discrimination is motivated by its interpretation as a measure of the relative likelihood for alternative probability densities. This is studied in a problem where a single sensor can be directed at any detection cell in the surveillance volume for each sample. Bayes rule is used to construct a recursive estimator for the cell target probabilities. The expected discrimination gain is predicted for each cell using its current target probability estimates. This gain is used to select the optimal cell for the next sample. For thresholded data, the expected discrimination gain depends on the threshold which is selected to maximize the gain for each sample. The expected discrimination gains can be maintained in a binary search tree structure for computational efficiency. The computational complexity of this algorithm is proportional to the height of the tree which is logarithmic in the number of detection cells. In a test case for a single 0 dB Gaussian target, the error rate for discrimination directed search was similar to the direct search result against a 6 dB target.

*Supported by the Institute for Mathematics and Its Applications under Air Force Office of Scientific Research contract AFOSR/F49620-94-1-0275

1 Introduction

The problem of sensor management is to determine how to select sensors, sensor modes and sensor search patterns to maximize their collective effectiveness. They may be located on different platforms, and the optimization measure may vary with mission requirements [Nash,Llinas, Papoli,Musick]. A typical application is to direct an electronically scanned aperture (ESA) radar [Blackman]. An ESA provides great flexibility in pointing and mode selection. For example, the beam can be redirected in a few microseconds, enabling targets to be illuminated at will. Different waveforms can be selected to provide more accurate range estimates or more accurate range-rates. The threshold setting can be changed to optimize search and track. Detection performance can be affected by using different pulse repetition intervals and radar frequencies. This is further complicated by the need to coordinate several sensors on a platform or group of platforms such as the use of a passive direction-finding sensor in conjunction with an ESA [Watson]. Another type of sensor management application arises when there is limited computational or communications band-width. An example of this type of problem is acoustic tracking using digital beam-forming. When low signal-to-noise targets are being tracked, adaptive beam-forming techniques may be required which are quite expensive computationally. In this situation, if it is not too computationaly expensive, a method for selecting which beams to form will provide improved performance for a fixed amount of computation.

One of the problems facing designers of sensor management systems is that there is little consensus on what mathematical quantities should be optimized. Much of the previous work in this area has focused on problems where the data are primarily kinematic

[Nash,Blackman,Schmaedeke], with little emphasis on the detection/classification element of the probelm. In these kinematic applications, a Kalman filter is used to predict how target state estimates will evolve in time based on estimated target dynamics, range, and signal/noise ratio. A typical application then solves for the update rate required to maintain the target position variance at some desired value, say, based on the trace of the position components of the covariance estimate. As target dynamics increase, then the required update rate also increases. A threshold in the dynamics can be used to switch the radar to a different mode such as a higher energy waveform when it is appropriate. This type of adaptive approach yields substantial improvement over non-adaptive sensor methods. However, it does not easily generalize to situations where detection, tracking and classification are being performed by the same set of sensors and must be simultaneously optimized.

It is often desirable to coordinate the collection of kinematic data with attribute information that can be used to identify a target. In order to simultaneously optimize conflicting objectives such as detection, tracking and classification, several authors have proposed the use of measures motivated by information theory [Hintz, Llinas]. This shifts the emphasis from optimizing parameter estimates for individual targets to globally optimizing the probability density estimates constructed by data fusion systems. This is advantageous because it can accommodate tradeoffs between sensors that measure different aspects of a scene. For example, consider the problem of coordinating the use of an ESA with an imaging sensor such as a Forward Looking Infrared (FLIR) camera. The FLIR provides azimuth and elevation information on the target. It can also be used to classify the target. Thus, FLIR provides both kinematic and attribute information to the user, but it is difficult to include the utility of the attributes in a measure such as the trace of the covariance matrix. A more elementary issue is that the covariance trace in the ESA example above only uses the position elements of the covariance matrix, so improvement in the velocity estimate of the target due to doppler information receives little weight in this scheme.

This paper is organized as follows. In the next section, use of discrimination gain for sensor management is motivated by its interpretation as a measure of the likelihood for different target probability densities. In Section 3, a detailed prediction of the discrimination gain is developed for the case of static targets confined to discrete detection cells. This is the principal result of the paper. The detection problem becomes one of classifying the state of each cell as either containing or not containing a target. The target detection/classification problem is obtained by extending the number of cell states to include different target types. One factor that simplifies the discrimination gain evaluation is that the probability estimates for the cells can be recursively estimated from the data. A second factor is that for each cell, the expected value of the discrimination obtained after an additional observation can be estimated without actually performing the observation. This is analogous to the use of the Riccati equation to predict the expected covariance after a target observation in a system using Kalman filters [Nash,Hintz]. Simulated results are presented in Section 4. For this example problem, the computational overhead per sample is only logarithmic in the number of sample cells. In order to provide a simple means of comparison with direct search, the probability of error is computed as a function of the average number of samples per cell for discrimination-based search against a 0 dB target and for direct search against 0 dB, 3 dB and 6 dB targets.

2 The Discrimination Function

Consider the problem of classifying objects that exist in S mutually exclusive and exhaustive discrete classes. Let $P[s]$ and $Q[s]$ be two probabilities for objects labeled by $s = 1, \ldots, S$. Then the discrimination of P with respect to Q is defined to be

$$D(P;Q) = \sum_{s=1}^{S} P[s] \ln(P[s]/Q[s]) \,. \qquad (1)$$

This is also refered to as the Kullback-Leibler information or cross-entropy[Kapur]. It can be shown that $D(P;Q)$ is non-negative and is 0 if and only if $P = Q$. As might be expected from its form as an average of the log-likelihood $\ln(P[s]/Q[s])$, the discrimination function is useful for characterizing the behavior of hypothesis testing problems [Blahut].

To understand the role of discrimination in sensor management, consider the problem of target detection/classification for a single detection cell. The true state of the cell, denoted $\tilde{s}$, is one of S states where $\tilde{s} = 1$ denotes no target, and $\tilde{s} = 2, \ldots, S$ denotes one of the possible target types. There is at most one target in the cell and based on prior information, the probability that a cell is in state s is given by the prior distribution $Q(s)$. Let $P(s)$ denote the estimated probability that the cell is in state s based on some measurements. Minimum information is given by the prior distribution, $P = Q$. Perfect information is represented by the probability

$$\tilde{P}(s) = \begin{cases} 1 & \text{if } s = \tilde{s}, \\ 0 & \text{otherwise.} \end{cases} \qquad (2)$$

and the discrimination of $\tilde{P}$ with respect to Q will be large.

The feature of discrimination that makes it useful in this context is the following. Suppose n objects are drawn at random from an infinite set of objects $\mathcal{O}$ labeled by a finite set of indices $s = 1, \ldots, S$. On any draw, the probability of selecting an object labeled by s is given by the prior probability $Q[s] = q_s$ (Q may be uniform, $q_s = 1/S$). The probability of drawing n_s occurances of object s is given by the multinomial distribution

$$Pr\{n_1, \ldots, n_S\} = \frac{n!}{n_1! \cdots n_S!} q_1^{n_1} q_2^{n_2} \cdots q_S^{n_S} \quad (3)$$

where $\sum_s n_s = n$. Defining $P[s] = n_s/n$, for fixed $P[s]$, it can be shown [Blahut] using Stirling's approximation that the probability of obtaining $\{n_1, \ldots, n_S\}$ obeys

$$Pr\{n_1, \ldots, n_S\} \sim e^{-nD(P;Q)+o(n)} \quad (4)$$

where D is the discrimination and $o(n)$ is a function that grows sublinearly with n.

An interpretation of Eq. (4) is that if the distribution P is very unlikely, then $D(P;Q)$ is large, and it has large information content relative to the a priori information. This is a plausible measure of the amount of information contained in P. In the detection/classification problem, cells whose estimates are far from the prior contain a great deal of information, while cells that are close to the prior contain little information. For the sensor manager, there is relatively little additional information to be derived from these well-characterized cells. For each sample, the expected discrimination gain is highest for the cells that currently have low discrimination values.

Discrimination can also be viewed informally as a distance measure between P and Q. However, it is important to note that discrimination does not enjoy all of the properties of a metric [Kapur]. In particular, it is not symmetric: $D(P;Q) \neq D(Q;P)$ in general. Also, it does not satisfy the triangle inequality. Interpreting the discrimination in this loose sense as a distance, maximizing discrimination serves to maximize the distance from the prior distribution. Heuristically, this is a useful strategy, assuming that the process of estimating P converges to $\tilde{P}$.

At first, it may seem that the use of maximum discrimination to guide sensor management is in conflict with Jaynes maximum entropy principal [Kapur]. This is not the case. In the context of this problem, the maximum entropy principal states that the most likely estimate of a probability distribution constructed from a collection of observations is the one that maximizes the entropy in the distribution or, equivalently, minimizes its information. This governs how the probability estimates are calculated, conditioned on the observations. On the other hand, sensor management provides guidance on the best observations to use for conditioning.

3 Discrimination Prediction

To understand how to compute expected information gain, consider the problem of detecting targets confined to discrete cells indexed by $c = 1, \ldots, C$. There is at most one target in each cell and the targets do not move. There are $S - 1$ target types. The state of each cell is one of $s = 1, \ldots, S$ values where $s = 1$ denotes no target in the cell and $s = 2, \ldots, S$ label the target types. The system is observed by a sensor that can be directed towards any single cell without pointing error. Only one cell is observed at a time. After one cell observation, the sensor can sample the same cell or any different cell. When the sensor is directed towards a cell it produces a measurement z which can be a discrete or continuous vector or set. The observation outcomes are independent and identically distributed (i.i.d.) random events. The conditional probabilities to obtain z given that cell c is in state s, $P_c[z|s]$ are known. Also, a priori probabilities for c to be in state s, $P_c[s]$ are known.

In this context, the detection/classification problem is to determine which state each cell is in, based on a set of observations. The outcome of the k-th observation of cell c is z_k (suppressing the c-index on the measurement). Let $\mathcal{Z}_K = (z_1, \ldots, z_K)$. After K observations, the posterior probability $P_c[s|\mathcal{Z}_K]$ that c is in state s conditioned on the observations is computed. For each cell c, the minimum-probability-of-error detector is obtained by choosing s to maximize $P_c[s|\mathcal{Z}_K]$ [Blahut].

Then sensor management problem is now to determine the optimal sensor parameters such as the threshold and the optimal order for searching the cells c. The heuristic developed here is to estimate the expected discrimination gain ΔD from the observation of any cell. This can be computed based on the current estimate of the cell state probabilities. The sensor manager then selects the cell that maximizes ΔD for each observation. This could be computationally intensive since in principal it involves summing over the entire observation set in $P_c[s|\mathcal{Z}_k]$ each time c is observed. However, it turns out that the evaluation of ΔD is relatively benign since $P_c[s|\mathcal{Z}_K]$ can be recursively evaluated from $P_c[s|\mathcal{Z}_{K-1}]$ and $P_c[z_K|s]$. Then ΔD can be computed from $P_c[s|\mathcal{Z}_K]$.

For each cell the probability that it is in state s, conditioned on K observations is computed using Bayes Rule as follows. (Since the expressions are the same for each cell, we suppress the cell subscript c here and in the sequel.) The total probability to observe $\mathcal{Z}_K$ is

$$P[\mathcal{Z}_K] = \sum_{s=1}^{S} P[\mathcal{Z}_K|s]P[s] \; . \quad (5)$$

Then since the data are i.i.d.,

$$P[\mathcal{Z}_K|s] = \prod_{k=1}^{K} P[z_k|s] \quad (6)$$

and the cell state probabilities are estimated using

$$P[s|\mathcal{Z}_K] = \frac{P[\mathcal{Z}_K|s]P[s]}{P[\mathcal{Z}_K]} \quad (7)$$

$$= \frac{P[\mathcal{Z}_K|s]P[s]}{\sum_{t=1}^{S} P[\mathcal{Z}_K|t]P[t]} \quad (8)$$

$$= P[s]\frac{\prod_{k=1}^{K} P[z_k|s]}{\sum_{t=1}^{S} P[t]\prod_{k=1}^{K} P[z_k|t]}\ , \quad (9)$$

where in the denominator summation, t indexes cell states.

After K observations, the discrimination of the estimated probability density with respect to the prior $P[s]$ is computed using the current values of $P[s|\mathcal{Z}_K]$ as

$$D[\mathcal{Z}_K] = \sum_{s=1}^{S} P[s|\mathcal{Z}_K]\ln\left(P[s|\mathcal{Z}_K]/P[s]\right)\ . \quad (10)$$

Observe that D depends on the observation set $\mathcal{Z}_K$ only through $P[s|\mathcal{Z}_K]$. Therefore, if $P[s|\mathcal{Z}_K]$ can be computed recursively, then so can D. $P[s|\mathcal{Z}_K]$ is obtained from Eq. (9) by substituting $P[s]\prod_{k=1}^{K-1} P[z_k|s] = P[s|\mathcal{Z}_{K-1}]P[\mathcal{Z}_{K-1}]$ in the numerator and denominator to yield the recursive expression

$$P[s|\mathcal{Z}_K] = \frac{P[z_K|s]P[s|\mathcal{Z}_{K-1}]}{\sum_{t=1}^{S} P[z_K|t]P[t|\mathcal{Z}_{K-1}]}\ . \quad (11)$$

To predict discrimination gain, note that after $K-1$ observations in a cell, the probability density for the K-th observation can be computed as

$$P[z_K|\mathcal{Z}_{K-1}] = \sum_{s=1}^{S} P[z_K|s]P[s|\mathcal{Z}_{K-1}]\ , \quad (12)$$

since the states s are mutually exclusive and exhaustive. The expected discrimination after one additional observation is obtained by summing over all possible measurement outcomes:

$$E[D|\mathcal{Z}_{K-1}] = \sum_{z_K} D[\mathcal{Z}_K]P[z_K|\mathcal{Z}_{K-1}]\ , \quad (13)$$

a function of the cell c and its observation set $\mathcal{Z}_{K-1}$. The expected discrimination gain for cell c is then

$$\Delta D[\mathcal{Z}_{K-1}, c] = E[D|\mathcal{Z}_{K-1}] - D[\mathcal{Z}_{K-1}]\ , \quad (14)$$

where the c-dependence of ΔD has been made explicit.

The discrimination-based sensor management strategy is to select the cell c that maximizes $\Delta D[\mathcal{Z}_K, c]$ (14). Here K is the current number of observations for the cell, which will in general be different for different cells. The algorithm can be made quite efficient if a list of the detection cells using $\Delta D[\mathcal{Z}_K, c]$ as the sort key is maintained as a binary search tree [Corman]. The list is maintained so that its right-most cell always maximizes the discrimination gain. After initialization, the algorithm can be written as:

1) Remove the node for the maximum gain cell $c^\star$ from the tree and observe the cell with the sensor.
2) Update $P_{c^\star}[s|\mathcal{Z}_K]$ with the new observation and evaluate $\Delta D[\mathcal{Z}_K, c^\star]$.
3) Insert $c^\star$ node in the tree, using its new value of ΔD to determine its location.
4) Return to 1).

To evaluate the computational complexity of this algorithm, observe that the cell state probabilities are independent so that computing the new discrimation gains does not depend on the the number of cells C. The parts of the algorithm that depend on C are deletion and insertion of the nodes, which, for a tree of height h, require $O(h)$ time. By using a balanced search-tree scheme such as a red-black tree [Corman], it is possible to guarantee that deletion and insertion can be accomplished in $O(\ln C)$ time so that the entire algorithm complexity is logarithmic in C.

4 An Example

First, a Bernoulli detection problem for thresholded data with single-observation detection probability $p_d = .8$ and single-observation false-alarm probability $p_{fa} = .5$ was studied. The expected number of targets in the surveillance volume is taken to be 1. The a priori target distribution is uniform in the volume so the prior probability for a target to be in each cell is $q = 1/C$. For this problem, the number of cell states S is 2 and the probabilities and discrimination are relatively simple to evaluate. For example, if the probability that a cell contains a target is p_t, then the discrimination obtained using Eq. (10) is

$$D = p_t\ln(p_t/q) + (1-p_t)\ln((1-p_t)/(1-q))\ . \quad (15)$$

Let $\tilde{p}_d$ be the average posterior probability computed using Eq. (11) that a target is present for the cell containing the target and $\tilde{p}_{fa}$ be the average posterior probability that a target is present for the cells that do not contain a target. A comparison was performed between the discrimination directed search and direct search for the case with $C = 100$ cells with the results averaged over several trials. In direct search, each cell in the volume is sampled sequentially and the number of samples is the same for each cell. For discrimination directed search, the result for 50 trials with 10 samples/cell was $\tilde{p}_d = .9$ and $\tilde{p}_{fa} = .001$ while for direct search, $\tilde{p}_d = .1$ and $\tilde{p}_{fa} = .01$ where obtained. This represents a substantial improvement in the accuracy of the probability estimate using discrimination-directed seach.

Next, the more complex problem of detecting a Gaussian target against a white noise background using thresholded data was examined. When no target is present, the expected signal strength is N_0. When a target is present, the expected signal strength is S_0+N_0. The optimal single sample detector for this problem is a North filter followed by Neyman-Pearson threshold test [Blahut]. With this procedure, the filter output λ is thresholded to produce a single measurement z which is 1 to indicate a threshold crossing, 0 otherwise. Using multiple measurements $\mathcal{Z}_k$, the probability that the target is present or absent is estimated for each cell using Eq. (11). For a North filter, when no target is present, the distribution for λ has variance $\sigma^2 = 2\ S_0/N_0$ and is

$$p_0(\lambda) = (2\pi\sigma^2)^{-1/2} \exp\left(-\lambda^2/2\sigma^2\right) . \qquad (16)$$

When the target is present, the distribution for λ is

$$p_1(\lambda) = (2\pi\sigma^2)^{-1/2} \exp\left(-(\lambda-\bar{\lambda})^2/2\sigma^2\right) \qquad (17)$$

where $\bar{\lambda} = 2\ S_0/N_0$. For a threshold setting τ, the single-sample detection probability is given by

$$p_d = \int_\tau^\infty p_1(\lambda) d\lambda \qquad (18)$$

while the single-sample false alarm probability is given by

$$p_{fa} = \int_\tau^\infty p_0(\lambda) d\lambda . \qquad (19)$$

There are now two subproblems that the sensor manager must solve for each cell sample: which cell should be sampled; and what should the threshold τ be? τ can be selected to maximize the expected discrimination gain ΔD for each sample. Figure 1 shows how ΔD varies with the threshold for several values of the estimated probability p_t that is a target in the cell when the target signal/noise ratio 0 dB. (The curves for ΔD are symmetric under the transformation $p_t \rightarrow 1-p_t$ and $\tau \rightarrow \bar{\lambda}-\tau$.) The expected discrimination gain is largest at $p_t = 1/2$ (the top curve) and the optimal threshold there is $\tau^\star = \bar{\lambda}/2$. Away from $p_t = 1/2$, the optimal threshold differs from $\bar{\lambda}/2$. This could be used to optimize the threshold for each cell, as a function of p_t. However, notice that as p_t moves away from 1/2, the τ-dependence of the gain decreases and that there is relatively little penalty incurred by using a fixed threshold of $\tau = \tau^\star$. This is the procedure followed here since it facilitates comparison with the direct search case. For $S_0/N_0 = 1$, $\tau = S_0/N_0$ yields $p_d = .76$ and $p_{fa} = .24$.

To compare results between direct search and discrimination-directed search, let c_0 denote the cell containing the target and $p_t(c,T)$ denote the estimated probability that the cell c contains a target based on data collected up to and including the time T. Then define an error probability

$$p_e(T) = Pr\{\arg\max_c\ p_t(c,T) \neq c_0\} \qquad (20)$$

Figure 2 shows results obtained for p_e versus the average number of samples per cell for dicrimination dircected search at 0 dB and direct search at 0 dB, 3 dB and 6 dB ($S_0/N_0 = 1, 2, 4$). Note that p_e falls much more quickly with the average number of samples using discrimination-directed search than direct search. Its performance at 0 dB is similar to the 6 dB direct-search case.

5 Discussion

The results presented in this paper suggest that optimization of discrimination gain is a useful criterion for managing sensors for detection and classification. The performance increase represented by Figure 2, obtained using discrimination-directed search for this model system, could be quite significant if similar increases hold for realistic systems. For typical active sensors the detection range for fixed signal/noise varies inversely as the 4-th power of the range. Therefore, a 6 dB gain in performance yields a 40% increase in detection range. For passive sensors, a 6 dB gain doubles the detection range. In operational scenarios, reducing time-to-decision is often the important issue. From Figure 2, if a system is operating at the 20% error rate against the 0 dB target, then discrimination-directed search achieves the decision point in only 3 samples/cell (300 total cell samples) while direct search requires 10 samples/cell (1000 total cell samples).

The method developed here for evaluating the expected gain is analogous to the covariance prediction step in a Kalman filter, with Eq. (12) used to compute the probability density for alternative measurement outcomes. The similarity in these computational processes suggests that discrimination gain can be used as a basis to manage sensors in situations that require simultaneous optimization of kinematics and attributes. In the Gaussian detection example studied here, a fixed threshold was used, independent of the value of the estimated target probability p_t in the cells. The potential performance gain from including the p_t-dependence of the threshold remains to be examined. Although the examples studied here are for thresholded data, thresholding is not required for using discrimination gain. It can also be be applied to non-thresholded data such as integrated acoustic energy or collections of attribute probabilities derived from electro-optical sensors.

Another issue for further study is that the basis for using discrimination is motivated heuristically here. An interesting open question is whether it can rigorously be shown that optimizing discrimination gain minimizes some probability of error such as Eq. (20). For example, Sanov's theorem [Blahut] states that in hypothesis testing problems, the asymptotic error probability for distinguishing between samples drawn from two distributions is characterized by the discrimination between them. This can be applied to

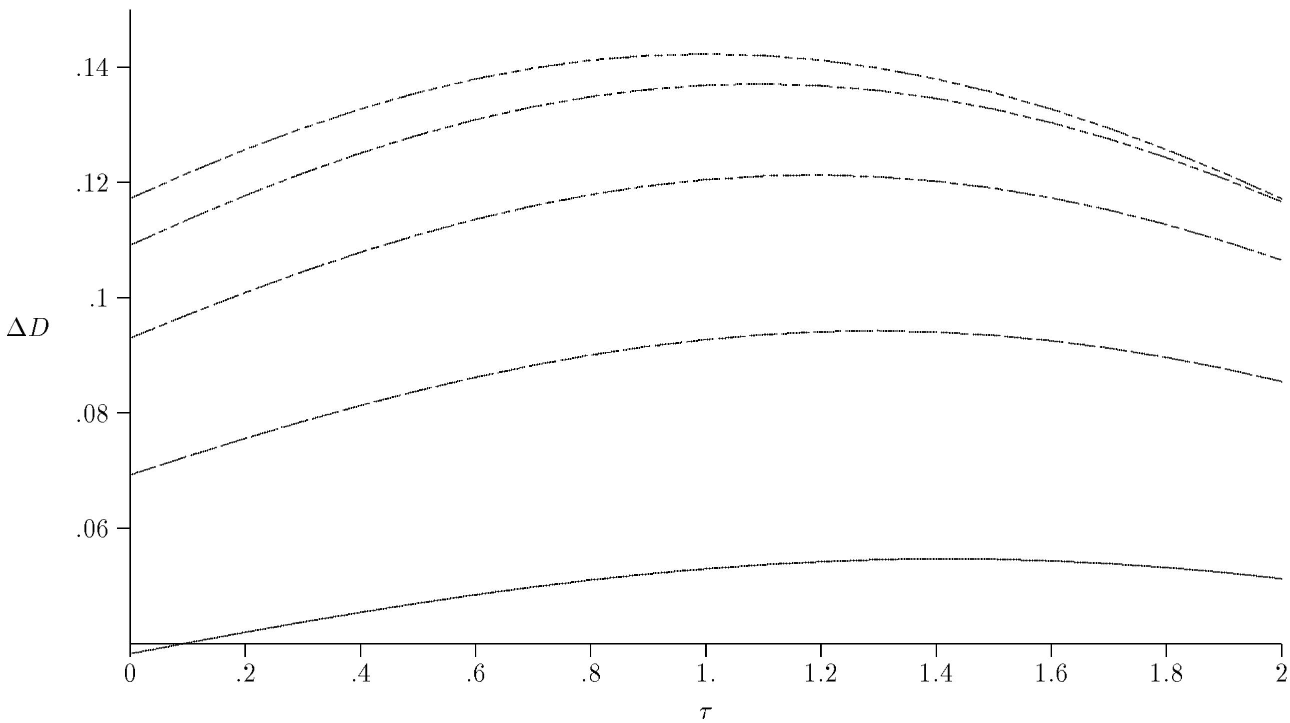

Figure 1: Expected discrimination gain ΔD vs. threshold τ for a 0 dB target. In increasing order, the curves are for target probability values of $p_t = .1, .2, \ldots, .5$.

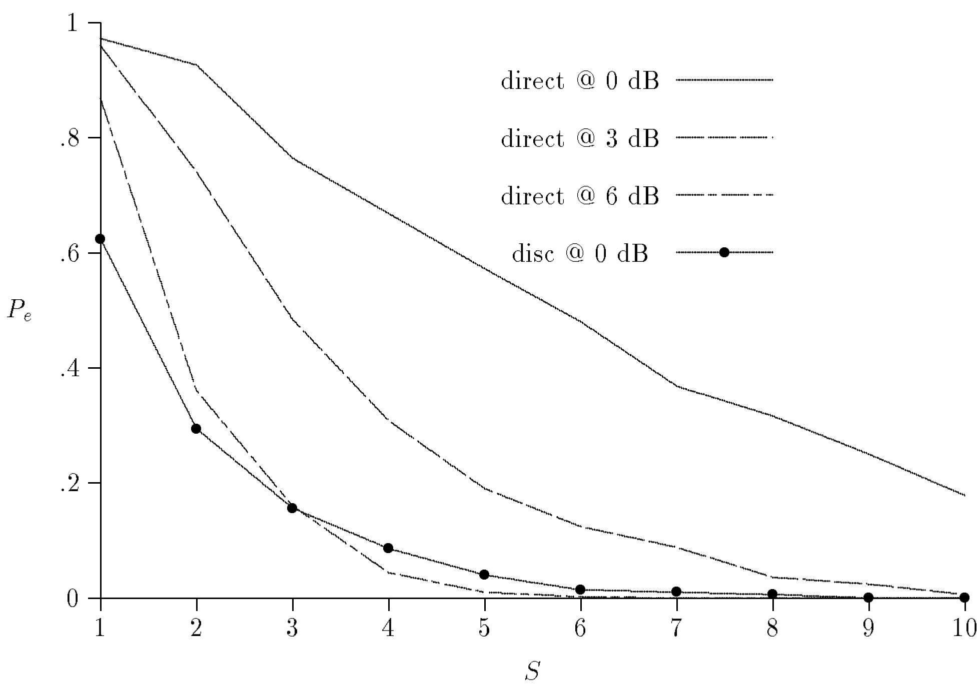

Figure 2: Error probability P_e as a function of average samples/cell S for discrimination-directed search (disc) and direct search. There are 100 cells in the in the surveillance volume. The curves are a 500-trial monte carlo average.

the sensor management context if it can be shown that maximizing the discrimination between the estimated density P and the prior Q results in asymptotically minimizing the discrimination between P and the true density $\tilde{P}$.

References

[Blackman] Samuel S. Blackman, *Multiple-Target Tracking with Radar Application*, Artech House, Norwood, MA, 1986.

[Blahut] R. E. Blahut, *Principles and Practice of Information Theory*,Addison-Wesley Pub. Comp., Reading, MA, 1987.

[Cormen] T. H. Cormen, C. E. Leiserson, and R. L. Rivest, *Introduction to Algorithms*, The MIT Press, Cambridge, MA, 1990.

[Hintz] K. J. Hintz and E. S. McVey, "Multi Process constrained estimation", IEEE Trans. Sys. Man Cyb., Vol. 21', No. 1, pp 237-244.

[Kapur] J. N. Kapur and H. K. Kesavan, *Entropy Optimzation Principles with Applications*, Academic Press, Inc., Boston, MA, 1992.

[Llinas] J. Llinas, K. J. Hintz and G. O. Beale, "Applications of Automatic Control Theory to Sensor Scheduling", Naval Air Warfare Center Technical Report NAWCADWAR-93002-50, 30 September, 1992.

[Musick] Stan Musick and Raj Malhotra, "Chasing the Elusive Sensor Manager", to appear, proceedings 1994 NAECON, Dayton, OH.

[Nash] J. M. Nash, "Optimal allocation of tracking resources", Proc. IEEE Conf. Dec. Cont., 1977, pp. 1177-1180.

[Popoli] Robert Popoli, "The Sensor Management Imperative", in Multitarget-Multisensor Tracking: Advanced Applications, Vol. II, Y. Bar-Shalom, Ed., Artech House, 1992.

[Watson] G. A. Watson and W. D. Blair, "IMM Algorithm for Tracking Targets that Maneuver Through Coordinated Turns", Proc. SPIE Signal and Data Processing of Small Targets, 1992, Orlando, FL, Apr. 1992, pp 236-247, "Revisit Control of a Phased-Array Radar for Tracking Maneuvering Targets when Supported by a Precision Electornic Support Measures Sensor", Proc. SPIE Signal and Data Processing of Small Targets, 1994, Orlando, FL, Apr. 1994, pp. 448-459.

[Schmaedeke] W. Schmaedeke, "Information-based Sensor Management", SPIE Proceedings Vol. 1955, Signal Processing, Sensor Fusion, and Target Recognition II, April, 1993. W. Schmaedeke and K. Kastella, "Information Theory and Sensor Management", Basic Research Group of the Joint Directors of Laboratories 1993 Symposium on Command and Control Research, June 28 - July 1, 1993.