Chapter 3

INFORMATION THEORETIC APPROACHES TO SENSOR MANAGEMENT

Alfred O. Hero

University of Michigan

Chris Kreucher

General Dynamics

1. Introduction

A good sensor management algorithm should only schedule those sensors that extract the highest quality information from the measurements. In recent years, several developers of new sensor management algorithms have used this compelling "folk theorem" as a fundamental design principle. This principle relies on tools of information theory to quantify and optimize the information collection capability of a sensor management algorithm. Any design methodology that uses such a guiding principle can be called an *information theoretic* approach to sensor management. This chapter reviews several of these approaches and explains the relevant information theory behind them.

A principal motivation behind information theoretic sensor management systems is that the system should be able to accommodate changing priorities in the mission of the sensing system, e.g., target detection, classification, or identification, as situational awareness evolves. Simply stated, the principal advantage of the information theoretic approach is that it simplifies system design by separating it into two independent tasks: information collection and risk/reward optimization. The sensor manager can therefore focus on the first task, optimizing information extraction, leaving the more complicated mission-specific details of risk-optimal tracking or classification to a downstream algo-

rithm. In other words, information theoretic sensor management approaches substitute a mission-independent surrogate reward function, the extracted information, for the mission-specific reward function of the standard POMDP approach described in Chapter 3.

Information theoretic approaches to selecting between different sources (sensors) of measurement have a long history that can be traced back to R. A. Fisher's theory of experimental design [87, 86], a field which remains very active to this day. In this setting one assumes a parametric model for the measurement generated by each sensor. Selecting a sensor is equivalent to selecting the likelihood function for estimating the value of one of the parameters. The mean (negative) curvature of the log likelihood function, called the Fisher information, is adopted as the reward. The optimal design is given by the choice of sensor that maximizes the Fisher information. When combined with a maximum likelihood parameter estimation procedure, the mission-specific part of the problem, this optimal design can be shown to asymptotically minimize the mean squared error as the number of measurements increases. Such an approach can be easily generalized to multiple parameters and to scheduling a sequence of sensor actions. However, it only has guaranteed optimality properties when a large amount of data is available to accurately estimate these parameters.

If one assumes a prior distribution on the parameters one obtains a Bayesian extension of Fisher's approach to optimal design. In the Bayesian setting each sensor's likelihood function induces a posterior density on the true value of the parameter. Posterior densities that are more concentrated about their mean values are naturally preferred since their lower spread translates to reduced parameter uncertainty. A monotone decreasing function of an uncertainty measure can be used as a reward function. The standard measure of uncertainty is the variance of the posterior, which is equal to the mean squared error (MSE) of the parameter estimator that minimizes MSE. A closely related reward is the Bayesian Fisher information, which is an upper bound on 1/MSE and has been proposed by [111]. These measures are only justifiable when minimum MSE makes sense, i.e., when the parameters are continuous and when the posterior is smooth (twice differentiable). An alternative measure of spread is the entropy of the posterior. Various measures of entropy can be defined for either discrete or continuous parameters and these will be discussed below.

In active sensor management where sensors are adaptively selected as measurements are made a more sensible strategy might be to maximize the rate of decrease of parameter uncertainty over time. This decrease in uncertainty is more commonly called the *information gain* and several relevant measures have been proposed including: the change in Shannon entropy of successive

posteriors [115, 114]; the Kullback divergence [211, 131, 168]; and the Rényi divergence [148] between successive posteriors. Shannon and Rényi divergence have strong theoretical justification as they are closely related to the rate at which the probability of decision error of the optimal algorithm decreases to zero.

The chapter is organized as follows. We first give relevant information theory background in Sec. 2, describing various information theoretic measures that have been proposed for quantifying information gains in sensing and processing systems. Then in Sec. 3 we describe methods of single stage optimal policy search driven by information theoretic measures. In Sec. 5 we show that in a precise theoretical sense that information gain can be interpreted as a proxy for any measure of performance. We then turn to illustrative examples including sensor selection for multi-target tracking applications in Sec. 6 and waveform selection for hyperspectral imaging applications in Sec. 7.

2. Background

The field of information theory was developed by Claude Shannon in the mid twentieth century [215] and has had an enormous impact on science in general, and in particular on communications, signal processing, and control. Shannon's information theory was the basis for breakthroughs in digital communications, data compression, and cryptography. A central result of information theory is the data processing theorem: physical processing of a signal cannot increase the amount of information it carries. Thus one of the main applications of information theory is the design of signal processing and communication systems that preserve the maximum amount of information about the signal. Among many other design tools, information theory has led to optimal and sub-optimal techniques for source coding, channel coding, waveform design, and adaptive sampling. The theory has also given general tools for assessing the fundamental limitations of different measurement systems, e.g., sensors, in achieving particular objectives such as detection, classification, or tracking. These fundamental limits can all be related to the amount of information gain associated with a specific measurement method and a specific class of signals. This leads to information gain methods of sensor management when applied to systems for which the user has the ability to choose among different types of sensors to detect an unknown signal. This topic will be discussed in the next section of this chapter.

Information theory provides a way of quantifying the amount of signal-related information that can be extracted from a measurement by its entropy, conditional entropy, or relative entropy. These information measures play cen-

tral roles in Shannon's theory of compression, encryption, and communication and they naturally arise as the primary components of information theoretic sensor management.

2.1 α-Entropy, α-Conditional Entropy, and α-Divergence

The reader is referred to the Appendix Section 1 for background on Shannon's entropy, conditional entropy, and divergence. Here we discuss the more general α-class of entropies and divergence used in the subsequent sections of this chapter.

Let $\mathbf{Y}$ be a measurement and $\mathbf{S}$ be a quantity of interest, e.g., the position of a target or the target identity (i.d.). We assume that $\mathbf{Y}$ and $\mathbf{S}$ are random variables with joint density $f_{Y,S}(y,s)$ and marginal densities f_Y and f_S, respectively. As discussed in the Appendix, Section 1, the Shannon entropy of $\mathbf{S}$, denoted $\mathcal{H}(S)$, quantifies uncertainty in the value of $\mathbf{S}$ before any measurement is made, called the prior uncertainty in $\mathbf{S}$. High values of $\mathcal{H}(S)$ imply high uncertainty about the value of $\mathbf{S}$. A more general definition than Shannon entropy is the alpha-entropy, introduced by I. Csiszár and A. Rényi [200]:

$$\mathcal{H}_\alpha(S) = \frac{1}{1-\alpha}\log \mathbb{E}[f_S^{\alpha-1}(\mathbf{S})] = \frac{1}{1-\alpha}\log \int f_S^\alpha(s)ds, \tag{3.1}$$

where we constrain α to $0 < \alpha < 1$. The alpha-entropy (4.1) reduces to the Shannon entropy (13.1) in the limit as α goes to one:

$$\mathcal{H}_1(S) \stackrel{\text{def}}{=} \lim_{\alpha\to 1}\mathcal{H}_\alpha(S) = -\int f_S(s)\log f_S(s)ds.$$

The conditional α-entropy of $\mathbf{S}$ given $\mathbf{Y}$ is the average α-entropy of the conditional density $f_{S|Y}$ and is related to the uncertainty of S after the measurement Y is made, called the posterior uncertainty. Similarly to (4.1) the conditional alpha-entropy is defined as

$$\begin{aligned}\mathcal{H}_\alpha(S|Y) &= \frac{1}{1-\alpha}\log \mathbb{E}[f_{S|Y}^{\alpha-1}(\mathbf{S}|\mathbf{Y})] \\ &= \frac{1}{1-\alpha}\log \int\int f_{S|Y}^\alpha(s|y)f_Y(y)dsdy,\end{aligned} \tag{3.2}$$

where again $0 < \alpha < 1$. A high quality measurement will yield a posterior density with low $\mathcal{H}_\alpha(S|Y)$ and given the choice among many possible sensors, one would prefer the sensor that yields a measurement that induces the lowest possible conditional entropy. This is the basis for entropy minimizing sensor management strategies.

We will also need the α-entropy of $\mathbf{S}$ conditioned on $\mathbf{Y}$ defined as

$$\mathcal{H}_\alpha(S|\mathbf{Y}) = \frac{1}{1-\alpha}\log \mathbb{E}[f_{S|Y}^{\alpha-1}(\mathbf{S}|\mathbf{Y})|\mathbf{Y}] = \frac{1}{1-\alpha}\log \int f_{S|Y}^{\alpha}(s|\mathbf{Y})ds.$$

As contrasted with the conditional α-entropy $\mathcal{H}_\alpha(S|Y)$, which is a non-random quantity, $\mathcal{H}_\alpha(S|\mathbf{Y})$ is a random variable depending on $\mathbf{Y}$. These two entropies are related by $\mathbb{E}[\mathcal{H}_\alpha(S|\mathbf{Y})] = \mathcal{H}_\alpha(S|Y)$.

A few comments about the role of the parameter $\alpha \in [0,1]$ are necessary. As compared to $f_{S|Y}$, $f_{S|Y}^{\alpha}$ is a function with reduced dynamic range. Reducing α tends to make $\mathcal{H}_\alpha(S|Y)$ more sensitive to the shape of the density $f_{S|Y}$ in regions of s where $f_{S|Y}(s|y) \ll 1$. As will be seen in Section 6, this behavior of $\mathcal{H}_\alpha(S|Y)$ can be used to justify different choices for α in measuring the reduction in posterior uncertainty due to taking different sensor actions.

As discussed in the Appendix Section 1, given two densities f, g of a random variable $\mathbf{S}$ the Kullback-Liebler divergence $\mathrm{KL}(f\|g)$ is a measure of similarity between them. Rényi's generalization, called the Rényi alpha-divergence, is

$$\mathcal{D}_\alpha(f\|g) = \frac{1}{\alpha-1}\log \int \left(\frac{f(s)}{g(s)}\right)^{\alpha} g(s)ds, \tag{3.3}$$

$0 < \alpha < 1$. There are other definitions of alpha-divergence that can be advantageous from a computational perspective, see e.g., [108], [6, Sec. 3.2] and [229]. These alpha-divergences are special cases of the information divergence, or f-divergence [71], and all have the same limiting form as $\alpha \to 1$. Taking this limit we obtain the Kullback Liebler divergence

$$\mathrm{KL}(f\|g) = \mathcal{D}_1(f\|g) \stackrel{\text{def}}{=} \lim_{\alpha\to 1} \mathcal{D}_\alpha(f\|g).$$

The simple multivariate Gaussian model arises frequently in applications. When f_0 and f_1 are multivariate Gaussian densities over the same domain with mean vectors μ_0, μ_1 and covariance matrices Λ_0, Λ_1, respectively [112]:

$$\begin{aligned}&\mathcal{D}_\alpha(f_1\|f_0) \qquad (3.4)\\ &= -\frac{1}{2(1-\alpha)}\ln\frac{|\Lambda_0|^{\alpha}|\Lambda_1|^{1-\alpha}}{|\alpha\Lambda_0+(1-\alpha)\Lambda_1|} + \frac{\alpha}{2}\Delta\mu^T(\alpha\Lambda_0+(1-\alpha)\Lambda_1)^{-1}\Delta\mu\end{aligned}$$

where $\Delta\mu = \mu_1 - \mu_0$ and $|A|$ denotes the determinant of square matrix A.

2.2 Relations Between Information Divergence and Risk

While defined independently of any specific mission objective, e.g., making the right decision concerning target presence, it is natural to expect a good

sensing system to exploit all the information available about the signal. Or, more simply put, one cannot expect to make accurate decisions without good quality information. Given a probability model for the sensed measurements and a specific task, this intuitive notion can be made mathematically rigorous.

2.2.1 Relation to Detection Probability of Error: the Chernoff Information.

Let $\mathbf{S}$ be an indicator function of some event, i.e. $\mathbf{S} \in \{0,1\}$ and $P(\mathbf{S}=1) = 1 - P(\mathbf{S}=0) = p$, for known parameter $p \in [0,1]$. Relevant events could be that a target is present in a particular cell of a scanning radar or that the clutter is of a given type. After observing the sensor output $\mathbf{Y}$ it is of interest to decide whether this event occurred or not and this can be formulated as testing between the hypotheses

$$\begin{aligned} H_0 &: \mathbf{S}=0 \\ H_1 &: \mathbf{S}=1. \end{aligned} \tag{3.5}$$

A test of H_0 vs. H_1 is a decision rule ϕ that maps $\mathbf{Y}$ onto $\{0,1\}$ where if $\phi(\mathbf{Y})=1$ the system decides H_1; otherwise it decides H_0. The 0-1 loss associated with ϕ is the indicator function $r_{0-1}(\mathbf{S},\phi(\mathbf{Y})) = \phi(\mathbf{Y})(1-\mathbf{S}) + (1-\phi(\mathbf{Y}))\mathbf{S}$. The optimal decision rule that minimizes the average probability of error $P_e = \mathbb{E}[r_{0-1}(\mathbf{S},\phi)]$ is the maximum a posteriori (MAP) detector which is a likelihood ratio threshold test ϕ^* where

$$\phi^*(\mathbf{Y}) = \begin{cases} 1, & \text{if } \frac{p(\mathbf{S}=1|\mathbf{Y})}{p(\mathbf{S}=0|\mathbf{Y})} > 1 \\ 0, & o.w. \end{cases}.$$

The average probability of error P_e^* of the MAP detector satisfies the Chernoff bound [69, Sec. 12.9]

$$P_e^* \geq \exp\left(\log \int f_{Y|S}^{\alpha^*}(y|1) f_{Y|S}^{1-\alpha^*}(y|\mathbf{S}=0)dy\right), \tag{3.6}$$

where

$$\alpha^* = \text{amin}_{0\leq\alpha\leq 1} \int f_{Y|S}^{\alpha}(y|1) f_{Y|S}^{1-\alpha}(y|0)dy.$$

The exponent in the Chernoff bound is identified as a scaled version of the α-divergence $D_{\alpha^*}\left(f_{Y|S}(\mathbf{Y}|1)\|f_{Y|S}(\mathbf{Y}|0)\right)$, called the *Chernoff exponent* or the *Chernoff information*, and it bounds the minimum log probability of error. For the case of n conditionally i.i.d. measurements $\mathbf{Y} = [\mathbf{Y}_1,\ldots,\mathbf{Y}_n]^T$ given $\mathbf{S}$, the Chernoff bound becomes tight as $n\to\infty$ in the sense that

$$-\lim_{n\to\infty}\frac{1}{n}\log P_e = (1-\alpha^*)D_{\alpha^*}\left(f_{Y|S}(\mathbf{Y}|1)\|f_{Y|S}(\mathbf{Y}|0)\right).$$

Thus, in this i.i.d. case, it can be concluded that the minimum probability of error converges to zero exponentially fast with rate exponent equal to the Chernoff information.

2.2.2 Relation to Estimator MSE: the Fisher Information.

Now assume that $\mathbf{S}$ is a scalar signal and consider the squared loss $(\mathbf{S} - \hat{\mathbf{S}})^2$ associated with an estimator $\hat{\mathbf{S}} = \hat{\mathbf{S}}(\mathbf{Y})$ based on measurements $\mathbf{Y}$. The corresponding risk $\mathbb{E}[(\mathbf{S} - \hat{\mathbf{S}})^2]$ is the estimator mean squared error (MSE), which is minimized by the conditional mean estimator $\hat{\mathbf{S}} = \mathbb{E}[\mathbf{S}|\mathbf{Y}]$

$$\min_{\hat{\mathbf{S}}} \mathbb{E}[(\mathbf{S} - \hat{\mathbf{S}})^2] = \mathbb{E}[(\mathbf{S} - \mathbb{E}[\mathbf{S}|\mathbf{Y}])^2].$$

The minimum MSE obeys the so-called Bayesian version of the Cramèr-Rao Bound (CRB) [234]:

$$\mathbb{E}[(\mathbf{S} - \mathbb{E}[\mathbf{S}|\mathbf{Y}])^2] \geq \frac{1}{\mathbb{E}[F(\mathbf{S})]},$$

or, more generally, the Bayesian CRB gives a lower bound on the MSE of any estimator of $\mathbf{S}$

$$\mathbb{E}[(\mathbf{S} - \hat{\mathbf{S}})^2] \geq \frac{1}{\mathbb{E}[F(\mathbf{S})]}, \tag{3.7}$$

where $F(s)$ is the conditional *Fisher information*

$$F(s) = \mathbb{E}\left[-\frac{\partial^2 f_{S|Y}(\mathbf{S}|\mathbf{Y})}{\partial \mathbf{S}^2} | \mathbf{S} = s \right]. \tag{3.8}$$

When the prior distribution of $\mathbf{S}$ is uniform over some open interval $F(s)$ reduces to the standard Fisher information for non-random signals

$$F(s) = \mathbb{E}\left[-\frac{\partial^2 f_{Y|S}(\mathbf{Y}|\mathbf{S})}{\partial \mathbf{S}^2} | \mathbf{S} = s \right]. \tag{3.9}$$

2.3 Fisher Information and Information Divergence

The Fisher information $F(s)$ can be viewed as a local approximation to the information divergence between the conditional densities $f_s \stackrel{\text{def}}{=} f_{Y|S}(y|s)$ and $f_{s+\Delta} = f_{Y|S}(y|s + \Delta)$ in the neighborhood of $\Delta = 0$. Specifically, let s be

a scalar parameter. A straightforward Taylor development of the α-divergence (4.3) gives

$$D_\alpha(f_s \| f_{s+\Delta}) = \frac{\alpha}{2} F(s)\Delta^2 + o(\Delta^2).$$

The quadratic term in Δ generalizes to $\frac{\alpha}{2}\underline{\Delta}^T \mathbf{F}(s)\underline{\Delta}$ in the case of a vector perturbation $\underline{\Delta}$ of a vector signal $\underline{s}$, where $\mathbf{F}(s)$ is the Fisher information matrix [6]. The Fisher information thus represents the curvature of the divergence in the neighborhood of a particular signal value s. This gives a useful interpretation for optimal sensor selection. Specializing to the weak signal detection problem, $H_0 : \mathbf{S} = 0$ vs. $H_1 : \mathbf{S} = \Delta$, we see that the sensor that minimizes the signal detector's error rate also minimizes the signal estimator's error rate. This equivalence breaks down when the signal is not weak, in which case there may exist no single sensor that is optimal for both detection and estimation tasks.

3. Information-Optimal Policy Search

At time $t = 0$, consider a sensor management system that direct the sensors to take one of M actions $a \in \mathcal{A}$, e.g., selecting a specific sensor modality, sensor pointing angle, or transmitted waveform. The decision to take action a is made only on the basis of past measurements $\mathbf{Y}_0$ and affects the distribution of the future measurement $\mathbf{Y}_1$. This decision rule is a mapping of $\mathbf{Y}_0$ to the action space $\mathcal{A}$ and is called a policy $\Pi(\mathbf{Y}_0)$. As explained in Chapter 3 of this book, the selection of an optimal policy involves the specification of the reward (or risk) associated with different actions. Recall that in the POMDP setting a policy generates a "(state, measurement, action)" sequence $\{(\mathbf{S}_0, \mathbf{Y}_0, a_0), (\mathbf{S}_1, \mathbf{Y}_1, a_1), \ldots\}$ and the quality of the policy is measured by the quality of the sequence of rewards $\{r(\mathbf{S}_1, a_0), r(\mathbf{S}_2, a_1), \ldots, \}$. In particular, with $\mathbb{E}[\mathbf{Y}|\mathbf{Y}]$ denoting the conditional expectation of $\mathbf{Y}$ given $\mathbf{Y}$, under broad assumptions the optimal policy that maximizes the discounted rewards is determined by applying Bellman's dynamic programming algorithm to the sequence of expected rewards $\{\mathbb{E}[r(\mathbf{S}_1, a_0)|\mathbf{Y}_0], \mathbb{E}[r(\mathbf{S}_2, a_1)|\mathbf{Y}_0, \mathbf{Y}_1], , \ldots, \}$.

For simplicity, in this section we will restrict our attention to single stage policies, i.e., myopic policies that seek only to maximize $\mathbb{E}[r(\mathbf{S}_1, a_0)|\mathbf{Y}_0]$. The basis for information gain approaches to sensor management is the observation that the expected reward depends on the action a_0 only through the information state $f_{S_1|Y_0,a_0}(s|Y_0, a_0)$:

$$\mathbb{E}[r(\mathbf{S}_1, a_0)|\mathbf{Y}_0] = \int r(s, a) f_{S_1|Y_0,a_0}(s|Y_0, a_0) ds.$$

The information state is the posterior density of the future state and its spread over state space $s \in \mathcal{S}$ is a measure of the uncertainty associated with predicting the future state $\mathbf{S}_1$ given the past measurement $\mathbf{Y}_0$ and the action a_0 dictated by the policy. Information gain strategies try to choose the policy that achieves maximum reduction in uncertainty of the future state. There are several information measures that can capture uncertainty reduction in the information state.

Perhaps the simplest measure of uncertainty reduction is the expected reduction in the variance of the optimal state estimator after an action a_0 is taken

$$\Delta U(a_0) = \mathbb{E}[(\mathbf{S}_1 - \mathbb{E}[\mathbf{S}_1|\mathbf{Y}_0])^2|\mathbf{Y}_0] - \mathbb{E}[(\mathbf{S}_1 - \mathbb{E}[\mathbf{S}_1|\mathbf{Y}_1, a_0])^2|\mathbf{Y}_0, a_0]$$

This measure is directly related to the expected reduction in the spread of the future information state $f_{S_1|Y_1,a_o}$ relative to that of the observed information state $f_{S_1|Y_0,a_o}$ due to action a_0. ΔU measures reduction in spread using the mean squared error norm squared, denoted $\|\mathbf{e}\| = \mathbb{E}[|\mathbf{e}|^2|\mathbf{Y}_0, a_0]$, where $\mathbf{e}$ is a prediction error. A wide variety of other types of norms can also be used, e.g., the absolute error norm $\|\mathbf{e}\| = \mathbb{E}\left[\,|\mathbf{e}|\,|\,\mathbf{Y}_0, a_0\right]$, to enhance robustness or otherwise emphasize/de-mphasize the tails of the posterior. Regardless of which norm is used, the optimum policy Π^* will achieve

$$\Pi^*(\mathbf{Y}_0) = \text{amin}_{a_0}\mathbb{E}\left[\,\|\mathbf{S}_1 - \mathbb{E}[\mathbf{S}_1|\mathbf{Y}_1, a_0]\|\|\,\mathbf{Y}_0, a_0\right]. \tag{3.10}$$

Another natural measure of uncertainty reduction is the expected change in entropy

$$\Delta U(a_0) = \mathcal{H}_\alpha(\mathbf{S}_1|\mathbf{Y}_0) - \mathbb{E}[\mathcal{H}_\alpha(\mathbf{S}_1|\mathbf{Y}_1)|\mathbf{Y}_0, a_0]$$

The optimal policy Π^* is obtained by replacing the norm $\|\mathbf{e}\|$ in (4.10) with the function $(\mathbf{e})^\alpha$ for $\alpha \in (0, 1)$ (Rényi entropy) or with $\log(\mathbf{e})$ (Shannon entropy) for $\alpha = 1$. The Shannon entropy version of this policy search method was used by Hintz [115] in solving sensor management problems for target tracking applications

The expected information divergence, called expected information gain, is another measure of uncertainty reduction:

$$\text{IG}_\alpha(a_0) = \mathbb{E}\left[\,\mathcal{D}_\alpha\left(f_{S_1|Y_1,a_0}(\mathbf{S}_1|\mathbf{Y}_1, a_0)\|f_{S_1|Y_0}(\mathbf{S}_1|\mathbf{Y}_0)\right)\Big|\,\mathbf{Y}_0, a_0\right], \tag{3.11}$$

which, for fixed $\mathbf{Y}_0$ and a_0, can be interpreted as the expectation over $\mathbf{Y}_1$ of the unobservable $\mathbf{Y}_1$-dependent information gain:

$$\begin{aligned}&\text{IG}_\alpha(\mathbf{Y}_1, a_0) \qquad\qquad (3.12)\\ &\qquad = \frac{1}{1-\alpha}\log\int\left(\frac{f_{S_1|Y_1,a_0}(S_1|\mathbf{Y}_1, a_0)}{f_{S_1|Y_0}(S_1|\mathbf{Y}_0)}\right)^\alpha f_{S_1|Y_0}(S_1|\mathbf{Y}_0)dS_1.\end{aligned}$$

When α approaches one the gain measure (4.11) reduces to the Kullback-Liebler divergence studied by Schmaedeke and Kastella [211], Kastella [131], Mahler [168], Zhao [260] and others in the context of sensor management.

When $f_{S_1|Y_0}$ is replaced by the marginal density f_{S_1}, as occurs when there is no Y_0 dependence, the α-divergence (4.12) reduces to the α-mutual information (α-MI) . As α converges to one the α-MI converges to the standard Shannon MI (see Section 1.4 of the Appendix). The Shannon MI has been applied to problems in pattern matching, registration, fusion and adaptive waveform design. The reader is referred to Chapter 11 Section 8 for more details on the latter application of MI.

Motivated by the sandwich bounds in Sec. 5, a definition of expected information gain that is more closely related to average risk can be defined

$$\begin{aligned} \Delta\overline{\mathrm{IG}}_\alpha(a_0) & \qquad (3.13) \\ &= \frac{1}{\alpha-1}\log \mathbb{E}\left[\left(\frac{f_{S_1|Y_1,a_0}(\mathbf{S}_1|\mathbf{Y}_1,a_0)}{f_{S_1|Y_0}(\mathbf{S}_1|\mathbf{Y}_0)}\right)^{\alpha}\middle|\,\mathbf{Y}_0,a_0\right], \end{aligned}$$

which can be expressed in terms of the unobservable $\mathbf{Y}_1$-dependent information gain (4.12) as:

$$\Delta\overline{\mathrm{IG}}_\alpha(a_0) = \frac{1}{\alpha-1}\log \mathbb{E}\left[e^{-(1-\alpha)\mathrm{IG}_\alpha(\mathbf{Y}_1,a_0)}\middle|\,\mathbf{Y}_0,a_0\right].$$

The choice of an appropriate value of α can be crucial to obtaining robust policies and this issue will be addressed in Section 6.

4. Information Gain Via Classification Reduction

A direct relation between optimal policies for sensor management and associated information divergence measures can be established using a recent result of Blatt and Hero [38] for reducing optimal policy search to an equivalent search for an optimal classifier. This strategy is called classification reduction of optimal policy search (CROPS) and leads to significantly more flexibility in finding approximations to optimal sensor management policies (see [39] for examples). The focus of this section is to show how CROPS leads us to a direct link between optimal policy search and information divergence measures.

The process of obtaining this relation is simple. An average reward maximizing policy is also a risk minimizing policy. A risk minimizing policy is equivalent to a classifier that minimizes a certain weighted probability of error for a related label classification problem. After a measure transformation

the weighted probability of error is equivalent to an unweighted probability of error, which is related to information divergence via the Chernoff bound. For simplicity of presentation, here we concentrate on binary action space and single stage policies.

Let the binary action space $\{A, B\}$ consist of the two actions A, B and define the associated rewards $\mathbf{r}_A = r(\mathbf{S}_1, A)$ and $\mathbf{r}_B = r(\mathbf{S}_1, B)$, respectively, when state $\mathbf{S}_1$ is observed after taking the specified action. Straightforward algebraic manipulations yield the following expression for the reward associated with policy Π [38]:

$$r(\mathbf{S}_1, \Pi(\mathbf{Y}_0)) = b - |\mathbf{r}_A - \mathbf{r}_B| I(\Pi(\mathbf{Y}_0) \neq \mathbf{C})$$

where $I(\mathcal{A})$ denotes the indicator of the event $\mathcal{A}$, $\mathbf{C}$ is the binary valued label $\mathbf{C} = \text{amax}_{a=A,B}\{\mathbf{r}_a\}$, and b is a constant independent of Π. With this expression the optimal single stage policy satisfies

$$\text{amax}_\Pi \mathbb{E}[r(\mathbf{S}_1, \Pi(\mathbf{Y}_0))|\mathbf{Y}_0] = \text{amin}_\Pi \tilde{\mathbb{E}}[I(\Pi(\mathbf{Y}_0) \neq \mathbf{C})|\mathbf{Y}_0], \quad (3.14)$$

where for any function g of the risk $\{\mathbf{r}_A, \mathbf{r}_B\}$, $\tilde{\mathbb{E}}(g(\mathbf{r}_A, \mathbf{r}_B)|\mathbf{Y}_0)$ denotes conditional expectation

$$\tilde{\mathbb{E}}(g(\mathbf{r}_A, \mathbf{r}_B)|\mathbf{Y}_0) = \int\int g(r_A, r_B) \tilde{f}_{r_A, r_B|Y_0}(r_A, r_B|\mathbf{Y}_0) dr_A dr_B,$$

and $\tilde{f}_{r_A, r_B|Y_0}$ is the "tilted" joint density function $f_{r_A, r_B}(r_A, r_B|Y_0)$ of r_A, r_B:

$$\tilde{f}_{r_A, r_B|Y_0}(r_A, r_B|\mathbf{Y}_0) = w(r_A, r_B) f_{r_A, r_B|Y_0}(r_A, r_B|\mathbf{Y}_0),$$

with weight factor

$$w(r_A, r_B) = \frac{|r_A - r_B|}{\mathbb{E}\left[|r_A - r_B| \mid \mathbf{Y}_0\right]}.$$

The relation (4.14) links the optimal risk minimizing policy Π to an optimal error-probability minimizing classifier of the random label $\mathbf{C}$ with posterior label probabilities: $P(\mathbf{C} = i|\mathbf{Y}_0) = \tilde{\mathbb{E}}[I(\mathbf{C} = i)|Y_0]$, $i = A, B$. Furthermore, by Chernoff's bound (4.6), the average probability of error of this optimal classifier has error exponent:

$$(1 - \alpha^*) \mathcal{D}_{\alpha^*}(f_A \| f_B), \quad (3.15)$$

where $f_A = f(Y_0|\mathbf{C} = A)$ and $f_B = f(Y_0|\mathbf{C} = B)$ are conditional densities of the measurement $\mathbf{Y}_0$ obtained by applying Bayes rule to $P(\mathbf{C} = A|\mathbf{Y}_0)$ and $P(\mathbf{C} = B|\mathbf{Y}_0)$, respectively. This provides a direct link between optimal sensor management and information divergence: the optimal policy is a Bayes optimal classifier whose probability of error decreases to zero at rate proportional to the information divergence (4.15).

5. A Near Universal Proxy

Consider a situation where a target is to be detected, tracked and identified using observations acquired sequentially according to a given sensor selection policy. In this situation one might look for a policy that is "universal" in the sense that the generated sensor sequence is optimal for all three tasks. A truly universal policy is not likely to exist since no single policy can be expected to simultaneously minimize target tracking MSE and target miss-classification probability, for example. Remarkably, policies that optimize information gain are near universal: they perform nearly as well as task-specific optimal policies for a wide range of tasks. In this sense the information gain can be considered as a proxy for performance for any of these tasks.

The fundamental role of information gain as a near universal proxy has been demonstrated both by simulation and by analysis in [147] and we summarize these results here. First we give a mathematical relation between marginalized alpha divergence and any task based performance measure. The key result is the following simple bound linking the expectation of a non-negative random variable to weighted divergence. Let $\mathbf{U}$ be an arbitrary r.v., let p and q be densities of U, and for any bounded non-negative (risk) function g define $\mathbb{E}_p[g(\mathbf{U})] = \int g(u)p(u)du$. Assume that q dominates p, i.e. $q(u) = 0$ implies $p(u) = 0$. Then, defining $w = \operatorname{ess\,inf} g(u)$ and $W = \operatorname{ess\,sup} g(u)$, Jensen's inequality immediately yields

$$w\mathbb{E}_q\left[\left(\frac{p(\mathbf{U})}{q(\mathbf{U})}\right)^{\alpha_1}\right] \leq \mathbb{E}_p[g(\mathbf{U})] \leq W\mathbb{E}_q\left[\left(\frac{p(\mathbf{U})}{q(\mathbf{U})}\right)^{\alpha_2}\right], \tag{3.16}$$

where $\alpha_1 \in [0, 1)$ and $\alpha_2 > 1$. Equality holds when $p = q$. This simple bound sandwiches any bounded risk function between two weighted alpha divergences.

Using the notation in Section 3 of this chapter, (4.16) immediately yields an inequality that sandwiches the predicted risk after taking an action a_0 by the expected information gain of form (4.13) with two different values of the Rényi divergence α parameter

$$we^{-(1-\alpha_1)\Delta\overline{\mathrm{IG}}_{\alpha_1}(a_0)} \leq \mathbb{E}[g(\mathbf{S}_1)|\mathbf{Y}_1, a_0] \leq We^{-(1-\alpha_2)\Delta\overline{\mathrm{IG}}_{\alpha_2}(a_0)}, \tag{3.17}$$

where $w = \inf_{y_0} \mathbb{E}[g(\mathbf{S}_1)|\mathbf{Y}_0 = y_0]$, $W = \sup_{y_0} \mathbb{E}[g(\mathbf{S}_1)|\mathbf{Y}_0 = y_0]$. This inequality is tight when α_1 and α_2 are close to one, and the conditional risk $\mathbb{E}[g(\mathbf{S}_1)|\mathbf{Y}_0]$ is only weakly dependent on the current measurement $\mathbf{Y}_0$.

In many cases of interest, one is only really concerned with estimation of a subset of the state variables $\mathbf{S}$. For example, for target tracking, the target state may be described by position velocity and acceleration but only the position of

the target is of interest. In such cases, the state can be partitioned into parameters of interest $\mathbf{U}$ and nuisance parameters $\mathbf{V}$, i.e., $\mathbf{S} = [\mathbf{U}, \mathbf{V}]$, and the risk function is constant with respect to $\mathbf{V}$, i.e., $g(\mathbf{S}) = g(\mathbf{U})$. According to (4.16), the appropriate sandwich inequality is modified from (4.17) by replacing $\mathbf{S}$ by $\mathbf{U}$. Specifically, the expected information gain (4.13) in the resultant bounds on the right and left of the inequality in (4.17) are replaced by expected IG expressions of the form

$$\begin{aligned} &\Delta\overline{\mathrm{IG}}_\alpha(a_0) \qquad (3.18)\\ &\quad = \frac{1}{\alpha-1}\log \mathbb{E}\left[\left(\frac{f_{U_1|Y_1,a_0}(\mathbf{U}_1|\mathbf{Y}_1,a_0)}{f_{U_1|Y_0}(\mathbf{U}_1|\mathbf{Y}_0)}\right)^\alpha \middle| \mathbf{Y}_0, a_0\right], \end{aligned}$$

which, as contrasted to the information gain (4.11), is expected IG between marginalized versions of the posterior densities, e.g.,

$$f_{U_1|Y_0}(\mathbf{U}_1|\mathbf{Y}_0) = \int f_{S_1|Y_0}(\mathbf{S}_1|\mathbf{Y}_0)d\mathbf{V}_1,$$

where $\mathbf{S}_1 = [\mathbf{U}_1, \mathbf{V}_1]$. We call the divergence (4.18) the marginalized information gain (MIG).

The sandwich inequality (4.17) is a theoretical result that suggests that the expected information gain (4.13) is a near universal proxy for arbitrary risk functions. Figure 4.1 quantitatively confirms this theoretical result for a simple single target tracking and identification example. In this simulation the target moves through a 100 $\times$ 100 cell grid according to a two dimensional Gauss-Markov diffusion process (see Sec. 6.1 for details). The moving target is one of 10 possible target types. At each time instant a sensor selects one of two modes, identification mode or tracking mode, and one of the 10,000 cells to query. In identification mode the sensor has higher sensitivity to the target type, e.g., a high spatial resolution imaging sensor, while in tracking mode the sensor has higher sensitivity to target motion, e.g., a moving target indicator (MTI) sensor. The output of these sensors was simply an integer-valued decision function taking values from 0 to 10. Output "0" denotes the "no target present" decision, output "Not 0" the "target present" decision, and output "k", $k \in \{1, \ldots, 10\}$ the "target is present and of class k" decision. The parameters (false alarm and miss probability, confusion matrix) of the sensor were selected to correspond to a realistic multi-function airborne surveillance system operating at 10dB SNR and to exhibit the tradeoff between tracking and identification performance.

The optimal target tracker and classifier are non-linear and intractable, as the measurement is non-Gaussian while state dynamics are Gaussian, and they were implemented using a particle filter as described in Chapter 6 of this book. Several policies for making sequential decisions on sensor mode and pointing direction were investigated: (1) a pure information gain (IG) policy that

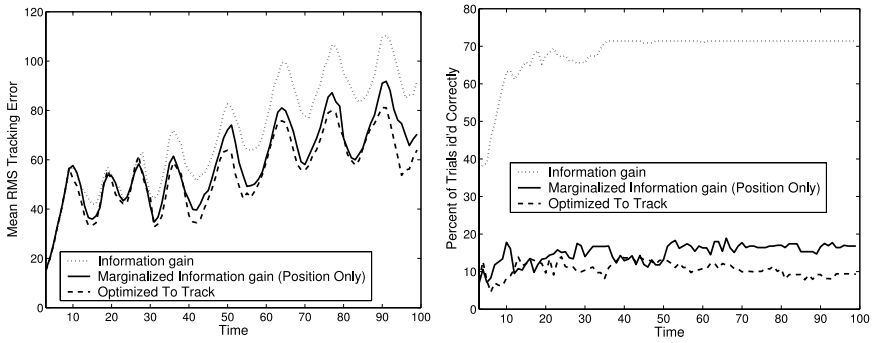

Figure 3.1. Left: comparisons between target position tracking rms error of a tracker that acquires data using sensor selection policies optimized under the information gain (IG), marginalized information gain (MIG), and tracking (rms) error reward functions. Right: same as left for target classification performance (Figure 1 from [147] - used with permission).

maximizes divergence between predicted posterior distributions of the four dimensional target state (position and velocity); (2) the marginalized IG (MIG) policy that maximizes the predicted divergence between posteriors of the two dimensional sub-state corresponding to position coordinate only; (3) a policy (rms) that minimizes predicted tracking mean squared error.

From the left panel of Fig. 4.1 it is evident that the IG optimized policy is not optimal for tracking the target and the performance of the optimal nonlinear tracker suffers due to the suboptimal IG policy. On the other hand, even though it is based only on information gain, the MIG optimized policy is nearly optimal for tracking as measured by the performance of the optimal tracker that uses MIG generated data. On the other hand, from the right panel of Fig. 4.1, we see that the IG policy does a much better job at classifying the target type. Thus, as predicted by the theoretical near-universality results in this section, the IG policy achieves a reasonable compromise between tracking and classification tasks.

6. Information Theoretic Sensor Management for Multitarget Tracking

In this section, we illustrate the efficacy of a specific information theoretic approach to sensor management which is based on the alpha divergence. We specialize to a multiple target tracking application consisting of estimating positions of a collection of moving ground targets using a sensor capable of interrogating portions of the surveillance region. Here the sensor management problem is one of selecting, on-line, the portion of the surveillance region to

be interrogated. This is done by computing the expected gain in information, as measured by the Rényi divergence, for each candidate action and taking the one with the maximum value.

6.1 The Model Multitarget Tracking Problem

The model problem is constructed to simulate a radar or EO platform, e.g., JSTARS, whose task is to track a collection of moving ground targets on the ground, assumed to be a plane. Specifically, there are ten ground targets moving in a $5km \times 5km$ surveillance region. Each target is described by its own 4 dimensional state vector $\mathbf{x}(t)$ corresponding to target position and velocity and assumed to follow the 2D diffusion model: $\dot{\mathbf{x}}_i(t) = \rho \mathbf{x}_i(t) + \mathbf{B}w_i(t)$, where ρ is the diffusion coefficient, $\mathbf{b} = [0, 0, \sigma_i, \sigma_i]^T$, and $w(t)$ is a standard (zero mean and unit variance) Gaussian white noise. The probability distribution of the target states is estimated on-line from sensor measurements via the JMPD model (see Chapter 5 of this book). This captures the uncertainty present in the estimate of the states of the targets. The true target trajectories come from a set of recorded data based on GPS measurements of vehicle positions over time collected as part of a battle training exercise at the Army's National Training Center.

The sensor simulates a moving target indicator (MTI) system in that at any time t_k, $k = 1, 2, \ldots,$, it lays a beam down on the ground that is one resolution cell (1 meter) wide and 10 resolution cells deep. The sensor is at a fixed location above the targets and there are no obscurations that would prevent a sensor from viewing a region in the surveillance area. The objective of the sensor manager is to select the specific 10 meter^2 MTI strip of ground to acquire. When measuring a cell, the imager returns either a 0 (no detection) or a 1 (a detection) which is governed by a probability of detection (P_d) and a per-cell false alarm rate (P_f). The signal to noise ratio (SNR) links these values together. In this illustrative example, we assume $P_d = 0.5$ and $P_f = P_d^{(1+SNR)}$, which is a model for a doppler radar using envelope detection (thresholded Rayleigh distributed returns). When there are T targets in the same cell the detection probability increases according to $P_d(T)$=$P_d^{\frac{1+SNR}{1+T\ SNR}}$; however the detector is not otherwise able to discriminate or spatially resolve the targets. Each time a beam is formed, a vector of measurements (a vector of zeros and ones corresponding to non-detections and detections) is returned, one measurement for each of the ten resolution cells.

6.2 Rényi Divergence for Sensor Scheduling

As explained above, the goal of the sensor is to choose which portion of the surveillance region to measure at each time step. This is accomplished by computing the value for each possible sensing action as measured by the Rényi alpha-divergence (4.3). In this multi-target tracking application, uncertainty about the multi-target state $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_T]^T$ and number T of targets at time conditioned on all the previous measurements $\mathbf{Y}^{k-1} = \{Y_1, \ldots, Y_{k-1}\}$ made up to and including time $k-1$ is captured by the JMPD (see Chapter 5 of this book) $f(\mathbf{X}^k, T^k|\mathbf{Y}^{k-1}, a_m)$. In this notation m (m = $1, \ldots, M$) will refer to the index of a possible sensing action $a_m \in \{a_1, \ldots, a_M\}$ under consideration, including but not limited to sensor mode selection and sensor beam positioning.

First the Rényi divergence between the current JMPD $f(\mathbf{X}^k, T^k|\mathbf{Y}^{k-1})$ and the updated JMPD $f(\mathbf{X}^k, T^k|\mathbf{Y}^k)$ must be computed. Therefore, we need

$$D_\alpha\left(f(\cdot|\mathbf{Y}^k)||f(\cdot|\mathbf{Y}^{k-1})\right) \tag{3.19}$$
$$= \frac{1}{\alpha-1}\log\int_{\mathbf{X}} f^\alpha(\mathbf{X}^k, T^k|\mathbf{Y}^k) f^{1-\alpha}(\mathbf{X}^k, T^k|\mathbf{Y}^{k-1}) d\mathbf{X}^k \ .$$

Using Bayes' rule, we can write

$$D_\alpha\left(f(\cdot|\mathbf{Y}^k)||f(\cdot|\mathbf{Y}^{k-1})\right) \tag{3.20}$$
$$= \frac{1}{\alpha-1}\log\frac{1}{f^\alpha(Y_k|\mathbf{Y}^{k-1}, a_m)}\int_{\mathbf{X}} f^\alpha(Y_k|\mathbf{X}^k, T^k, a_m) f(\mathbf{X}^k, T^k|\mathbf{Y}^{k-1}) d\mathbf{X}^k \ .$$

Our aim is to choose the sensing action to take *before actually receiving* the measurement Y_k. Specifically, we would like to choose the action that makes the divergence between the current density and the density after a new measurement as large as possible. This indicates that the sensing action has maximally increased the information content of the measurement updated density, $f(\mathbf{X}^k, T^k|\mathbf{Y}^k)$, with respect to the density before a measurement was made, $f(\mathbf{X}^k, T^k|\mathbf{Y}^{k-1})$. However, we cannot choose the action that maximizes the divergence as we do not know the outcome of the action before taking it. As an alternative, as explained in Section 5, we calculate the expected value of (4.20) for each of the M possible sensing actions and choose to take the action that maximizes the expectation. Given past measurements $\mathbf{Y}^{k-1}$, the expected value of (4.20) may be written as an integral over all possible measurement outcomes $\mathbf{Y}_k = y$ when performing sensing action a_m as

$$\mathbb{E}[D_\alpha|\mathbf{Y}^{k-1}] = \int f(y|\mathbf{Z}^{k-1}, a_m) D_\alpha\left(f(\cdot|\mathbf{Y}^k)||f(\cdot|\mathbf{Y}^{k-1})\right) dy \ . \tag{3.21}$$

In analogy to (4.13) we refer to this quantity as the expected information gain associated with sensor action a_m.

6.3 Multitarget Tracking Experiment

The empirical study shown in Figure 4.2 shows the benefit of the information theoretic sensor management method. In this figure, we compare the performance of the information theoretic method where sensing locations are chosen based on expected information gain, a periodic method where the sensor is sequentially scanned through the region, and two heuristic methods based on interrogating regions where the targets are most likely to be given the kinematic model and the estimated positions and velocities at the previous time step (see [150] for more detailed explanation). We compare the performance by looking at root-mean-square (rms) error versus number of sensor resources available ("looks"). All tests use the true SNR ($= 2$) and are initialized with the true number of targets and the true target positions.

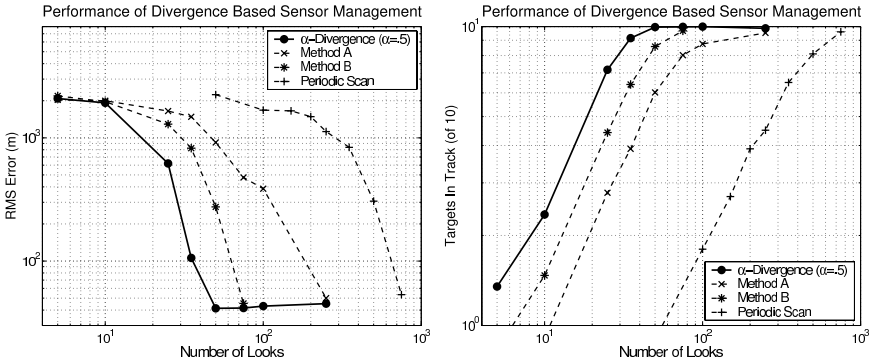

Figure 3.2. A comparison of the information-based method to periodic scan and two other methods. The performance is measured in terms of the (median) rms error versus number of looks and the (average) number of targets in track. The α-divergence strategy out performs the other strategies, and at 35 looks performs similarly to non-managed with 750 looks. (Left panel is Figure 6 of [150] - used with permission)

6.4 On the Choice of α

The Rényi divergence has been used in many applications, including content-based image retrieval, image georegistration, and target detection [113, 112]. These studies provide guidance as to the optimal choice of α.

In the georegistration problem [112] it was empirically determined that the value of α leading to highest resolution clusters around either $\alpha = 1$ or $\alpha = 0.5$ corresponding to the KL divergence and the Hellinger affinity respectively. The determining factor appears to be the degree of to which the two densities under consideration are different. If the densities are very similar then the indexing performance of the Hellinger affinity distance ($\alpha = 0.5$) was observed to be better than that of the KL divergence ($\alpha = 1$). Furthermore, the asymptotic analysis of [112] shows that $\alpha = .5$ provides maximum discrimination between two similar densities. This value of α provides a weighting which stresses the tails, or the minor differences, between two distributions. In target tracking applications with low measurement SNR and slow target dynamics, with respect to the sampling rate, the future posterior density can be expected to be only a small perturbation on the current posterior density, justifying the choice of $\alpha = 0.5$.

Figure 4.3 gives an empirical comparison of the performance under different values of α. All tests use $P_d = 0.5$, $SNR = 2$, and $P_f = P_d^{(1+SNR)}$. We find that $\alpha = 0.5$ performs best here as it does not lose track on any of the 10 targets during any of the 50 simulation runs. Both cases of $\alpha \approx 1$ and $\alpha = 0.1$ case cause frequent loss of track of targets.

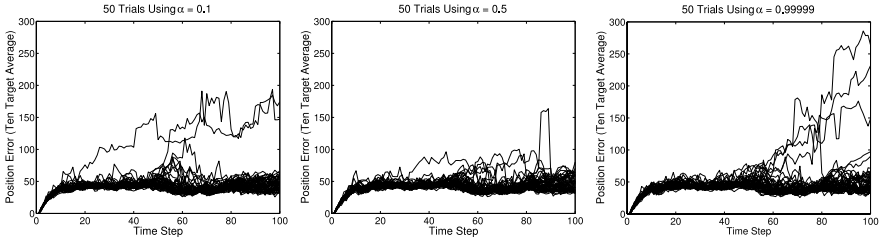

Figure 3.3. A comparison of sensor management performance under different values of α. On simulations involving ten real targets, $alpha = 0.5$ leads to the best tracking performance. (Figure 5 of [150] - used with permission)

6.5 Sensitivity to Model Mismatch

Here we present empirical results regarding the performance of the algorithm under model mismatch. Computation of the JMPD and information gain requires accurate models of target kinematics and the sensor. In practice, these models may not be well known. Figure 6.5 shows the effect of mismatch between the assumed target kinematic model and the true model. Specifically, Fig. 6.5 shows the sensitivity to mismatch in the assumed diffusion coefficient ρ and the noise variance $\sigma_i = \sigma$, equivalently, the sensor SNR, relative to their

true values used to generate the data. The vertical axis of the graph shows the true coefficient of diffusion of the target under surveillance. The horizontal axis shows the mismatch between the filter estimate of kinematics and the true kinematics (matched = 1). The color scheme shows the relative degradation in performance present under mismatch (< 1 implies poorer performance than the matched filer). The graph in Fig. 6.5 shows (a) how a poor estimate of the kinematic model effects performance of the algorithms, and (b) how a poor estimate of the sensor SNR effects the algorithm. In both cases, we find that the information gain method is remarkably robust to model mismatch, with a graceful degradation in performance as the mismatch increases.

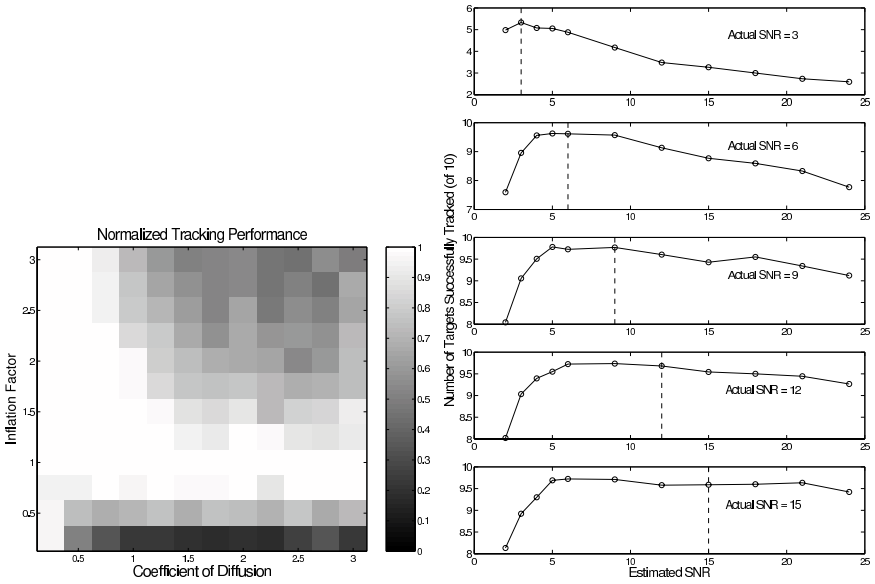

Figure 3.4. Left: Performance degradation when the kinematic model is mismatched. Performance degrades gradually particularly for high SNR targets.

6.6 Information Gain vs Entropy Reduction

Another information theoretic method that is very closely related to maximizing the Rényi divergence is maximization of the expected change in Rényi entropy. This method proceeds in a manner nearly identical to that outlined above, with the exception that the metric to be maximized is the expected change in entropy rather than the expected divergence. It might be expected that choosing sensor actions to maximize the decrease in entropy may be better than trying to maximize the divergence since entropy is a more direct measure

of concentration of the posterior density. On the other hand, unlike the divergence, the Rényi entropy difference is not sensitive to the dissimilarity of the old and new posterior densities and does not have the strong theoretical justification provided by the error exponent (4.15) of Sec. 3. Generally we have found that if there is a limited amount of model mismatch, maximizing Rényi either divergence or entropy difference gives equivalent sensor scheduling results. For example, for this multi-target tracking application, an empirical study (Figure 6.6) shows that the two methods yield very similar tracking performance.

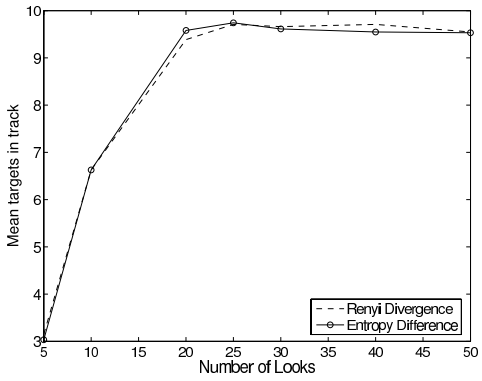

Figure 3.5. Performance, in terms of the number of targets successfully tracked, when using a sensor management strategy predicated on either the Rényi divergence of the change in entropy.

7. Terrain Classification in Hyperspectral Satellite Imagery

A common problem in passive or active radar sensing is the waveform selection problem. For more detailed discussion of active radar waveform design and selection the reader is referred to Chapter 11 of this book. Different waveforms have different capabilities depending on the target type, the clutter type and the propagation characteristics of the medium. For example, accurate discrimination between some target and clutter scenarios may require transmission of the full set of available waveforms while in other scenarios one may get by with only a few waveforms. In many situations the transmitted energy or the processed energy are a limited resource. Thus, if there is negligible loss in performance, reduction of the average number of waveforms transmitted is desirable. the problem of selection of an optimal subset of the available waveforms is relevant. This is an instance of an optimal resource allocation problem for active radar.

This section describes a passive hyperspectral radar satellite-to-ground imaging application for which the available waveforms are identified with different spectral bands. To illustrate the methods discussed in this chapter we use the Landsat satellite image dataset [222]. This dataset consists of a number of examples of ground imagery and is divided into training data (4435 examples) and test data (2000 examples) segments. The ground consists of six different classes of earth : "red soil", "cotton crop", "grey soil", "damp grey soil", "soil with vegetation subtle", and "very damp grey soil". For each patch of the ground, the database contains a measurement in each of four spectral bands, ranging from visible to infra-red. Each band measures emissivity at a particular wavelength in each of the 6435 pixelated 3×3 spatial regions. Thus the full four bands give a 36 dimensional real valued feature vector. Furthermore, the database comes with ground truth labels which associates each example with the type of earth corresponding to the central pixel of each 3×3 ground patch.

We consider the following model problem. Assume that the sensor is only allowed to measure a number p of the four spectral bands where $p < 4$. When we come upon an unidentified type of earth the objective is to choose the collection of bands that will give the most information about the class. Thus here the state variable x is the unknown class of the terrain type. We assume that at initialization only prior information extracted from the data in the training set is available. Specifically, denote by $p(x)$ the prior probability mass function for class x, $x \in \{1, 2, 3, 4, 5, 7\}$. The set of relative frequencies of class labels in the training database implies that $p(x = 1) = 0.24$, $p(x = 2) = 0.11$, $p(x = 3) = 0.22$, $p(x = 4) = 0.09$, $p(x = 5) = 0.11$, and $p(x = 7) = 0.23$. Denote by $f(\mathbf{Y}|x = c)$ the multivariate probability density function of the (9p)-dimensional measurement $\mathbf{Y} = [Y_{i_1}, \ldots, Y_{i_p}]^T$ of the 3×3 terrain patch when selecting the combination $B = \{i_1, \ldots, i_p\}$ of spectral bands and when the class label of the terrain is $x = c$.

7.1 Optimal Waveform Selection

Here we explore optimal off-line waveform selection based on maximizing information gain as compared to waveform selection based on minimizing miss-classification error P_e. The objective in off-line waveform selection is to use the training data to specify a single best subset of p waveform bands that entails minimal loss in performance relative to using all 4 waveform bands. Off-line waveform selection is to be contrasted to on-line approaches that account for the effect on future measurements of waveform selection based on current measurements $\mathbf{Y}_0$. We do not explore the on-line version of this problem here. Online waveform design for this hyperspectral imaging example is reported in [39] where the classification reduction of optimal policy search

(CROPS) methodology of Sec. 4 is implemented to optimally schedule measurements to maximize terrain classification performance.

7.1.1 Terrain Miss-Classification Error. The miss-classification probability of error P_e of a good classifier is a task specific measure of performance that we use as a benchmark for studying the information gain measure. For the Landsat dataset the k-nearest neighbor (kNN) classifier with $k = 5$ has been shown to perform significantly better than other more complex classifiers [107] when all four of the spectral bands are available. The kNN classifier assigns a class label to a test vector $\mathbf{y}$ by taking a majority vote among the labels of the k closest points to $\mathbf{y}$ in the training set. The kNN classifier is non-parametric, i.e., it does not require a model for the likelihood function $\{p(\mathbf{z}_B|x = c)\}_{c=1}^{6}$. However, unlike model-based classifiers that only require estimated parameter values obtained from the training set, the full training set is required to implement the kNN classifier. The kNN classifier with $k = 5$ was implemented for all possible combinations of 4 bands to produce the results below (Tables 7.1.3-4.2).

7.1.2 Terrain Information Gain. To compute the information gain we assume a multivariate Gaussian model for the likelihood $f(Y_k|X_k)$ and infer its parameters, i.e., the mean and covariance, from the training data. Since x is discrete valued the Rényi divergence using the combination of bands B is simply expressed:

$$\langle D_\alpha \rangle_B = \int_{\mathbf{z}_B} p(\mathbf{z}_B) \frac{1}{\alpha - 1} \log \sum_{x=1}^{6} p(x)^\alpha p(x|\mathbf{z}_B)^{1-\alpha} d\mathbf{z}_B \ , \tag{3.22}$$

where

$$p(x|\mathbf{z}_B) = \frac{p(x|\emptyset) p(\mathbf{z}_B|x)}{p(\mathbf{z}_B)} \ . \tag{3.23}$$

All of the terms required to compute these integrals are estimated by empirical moments extracted from the training data. The integral must be evaluated numerically as, to the best of our knowledge, there is no closed form.

7.1.3 Experimental Results. For the supplied set of Landsat training and test data we find the expected gain in information and miss-classification error P_e as indicated in Tables 7.1.3, 7.1.3 and 4.2.

These numbers are all to be compared to the "benchmark" values of information gain, 1.30, and the misclassification probability, 0.96, when all 4

Single Band	Mean Info Gain	Pe(kNN)
1	0.67	0.379
2	0.67	0.347
3	0.45	0.483
4	0.75	0.376

Table 3.1. Expected gain in information and Pe of kNN classifier when only a single band can be used. The worst band, band 3, provides the minimum expected gain in information and also yields the largest P_e. Interestingly, the single bands producing maximum information gain (band 4) and minimum P_e (band 2) are different.

Band Pair	Mean Info Gain	Pe(kNN)
1,2	0.98	0.131
1,3	0.93	0.134
1,4	1.10	0.130
2,3	0.90	0.142
2,4	1.08	0.127
3,4	0.95	0.237

Table 3.2. Expected gain in information and P_e of kNN classifier when a pair of bands can be used. The band pair (1,4) provides the maximum expected gain in information followed closely by the band pair (2,4), which is the minimum P_e band pair.

Band Triple	Mean Info Gain	Pe(kNN)
2,3,4	1.17	0.127
1,3,4	1.20	0.112
1,2,4	1.25	0.097
1,2,3	1.12	0.103

Table 3.3. Expected gain in information and P_e of kNN classifier when only three bands can be used. Omitting band 3 results in the highest expected information gain and lowest P_e.

spectral bands are available. Some comments on these results will be useful. First, if one had to throw out a single band, use of bands 1,2,4 entails only a very minor degradation in P_e from the benchmark and the best band to eliminate (band 3) is correctly predicted by the information gain. Second, the small discrepancies between the ranking of bands by P_e and information criteria can be explained by several factors: 1) the kNN is non-parametric while the information gain imposes a Gaussian assumption on the measurements; 2) the information gain is only related to P_e indirectly, through the Chernoff bound; 3) the kNN classifier is not optimal for this dataset - indeed recent results show that a 10% to 20% decrease in P_e is achievable using dimensionality reduction

techniques [198, 107]. Finally, as these results measure average performance they are dependent on the prior class probabilities, which have been determined from the relative frequencies of class labels in the training set. For a different set of priors the results could change significantly.

8. Conclusion and Perspectives

The use of information theoretic measures for sensor management and waveform selection has several advantages over task-specific criteria. Foremost among these is that, as they are defined independently of any estimation of classification algorithm, information theoretic measures decouple the problem of sensor selection from algorithm design. This allows the designer to hedge on the end-task and go after designing the sensor management system to optimize a more generic criterion such as information gain or Fisher information. In this sense, information measures are similar to generic performance measures such as front end signal-to-noise-ratio (SNR), instrument sensitivity, or resolution. Furthermore, as we have shown here, information gain can be interpreted as a near universal proxy for any performance measure. On the other hand, information measures are more difficult to compute in general situations since, unlike SNR, they may involve evaluating difficult non-analytical integrals of functions of the measurement density. There remain many open problems in the area of information theoretic sensor management, the foremost being that no general real-time information theory yet exists for systems integrating sensing, communication, and control.

A prerequisite to implementation of information theoretic objective functions in sensor management is the availability of accurate estimates of the posterior density (belief state) of the state given the measurements. In the next chapter a general joint particle filtering approximation is introduced for constructing good approximations for the difficult problem of multiple target tracking with target birth/death and possibly non-linear target state and measurement equations. In Chapter 6 this approximation will be combined with the information gain developed in this chapter to perform sensor management for multiple target trackers. Information theoretic measures are also applied in Chapter 11 for adaptive radar waveform design.