Proceedings of the 2007 American Control Conference
Marriott Marquis Hotel at Times Square
New York City, USA, July 11-13, 2007

Comparison of Information-Theoretic Objective Functions for Decision Support in Sensor Systems

Chenghui Cai and Silvia Ferrari

Abstract— **Information-driven sensor management aims at making optimal decisions regarding the sensor type, mode and configuration in view of the sensing objectives. In this paper, an approach is developed for computing two information-theoretic functions, expected discrimination gain and expected entropy reduction, to optimize target classification accuracy based on multiple and heterogeneous sensors fusion. The measurement process is modeled by means of Bayesian networks (BNs). The two objective functions utilize the BN models to represent the expected effectiveness of the sensors search sequence. New theoretic solutions are presented and implemented for computing the objective functions efficiently, based on the BN factorization of the underlying joint probability distributions. Dempster-Shafer fusion rule is embedded in the computations in order to account for the complementarity of multiple, heterogeneous sensor measurements. The efficiency of the two objective functions is demonstrated and compared using a landmine detection and classification application.**

I. Introduction

THE problem of information-driven sensor planning and management for target classification consists of optimally deciding the sensor type and mode that maximize the expected information profit. The sensor information profit is defined as the expected value of the information obtained through the sensor measurements, minus the cost associated with the use of the sensor and related resources, such as, its platform. The main philosophy behind this approach is to base the decision for sensor planning and platform navigation on dynamic sensor measurements that become available over time and whose outcome depends on the decision variables. For many sensor surveillance systems involving multiple and heterogeneous components or agents, the value of sensor measurements can be expressed as an information-theoretic objective function. Then, the measurements can be viewed as a feedback to the sensor manager (or controller), and can be used to make optimal decisions about measurement sequence and sensor parameters. Ultimately, the solutions must optimize classification accuracy, probability of detection, and minimize the probability of false alarms.

The use of information-theoretic objective functions for sensor management has been proposed by several authors. Schmaedeke used a discrimination gain technique to solve a multisensor-multitarget assignment problem [1]. Kastella managed agile sensors to optimize detection and classification based on discrimination gain [2]. Zhao investigated information objective functions such as entropy and Mahalanobis distance measure for sensor collaboration applications [3]. However, little work has been done to compare these objective functions and analyze their performance across distinct sensor applications, such as, feature estimation and target classification. In this paper, a BN framework is developed for computing the discrimination gain and entropy reduction in multiple and heterogeneous sensor systems.

Two common applications of multiple sensor systems are the classification of the target features from fused sensor measurements, referred to as feature inference, and target classification. The problems of feature inference for a Gaussian target are provided in [2]. In this paper, the theoretic solutions for non-Gaussian distributions are derived for both feature inference and target classification.

When multiple heterogeneous sensors are employed, their complementarity and performance relative to the environmental conditions are exploited through fusion. Dempster-Shafer (D-S) fusion technique has been shown to be very effective for performing feature inference and target classification based on multiple and heterogeneous sensor measurements [4-7]. A novel contribution of this paper is that the D-S fusion rule is embedded in the computations of the information objective functions to evaluate the expected benefit of obtaining sensor information that will be fused *a posteriori*. Also, the Bayesian network (BN) sensor modeling presented in [8] is used in order to obtain a methodology that can be generalized to any measurement process, regardless of the form of the underlying probability distributions.

The paper is organized as follows. In Section II, the discrimination gain and entropy reduction are introduced. In Section III, the computation of these objective functions is presented for the feature-inference and target-classification cases, and D-S fusion is incorporated in the BN classification frame. The demining application is presented and demonstrated in Section IV.

II. Background

A. Bayesian Network Modeling of Sensor Measurements

A Bayesian network (BN) model is a directed acyclic graph (DAG) [8] comprised of a set of nodes representing variables, and a set of directed arcs connecting the nodes. In this paper, capital letters denote sets of variables, lowercase letters denote variables, and subscripts in lowercase letters denote the possible *states* of the variables. In BN models, Bayes' rule of inference is utilized together with graphical manipulations to compute the posterior probability distribu-

Manuscript received September 24, 2006.
C. Cai is a graduate student of Mechanical Engineering at Duke University, Durham, NC, 27708, USA cc88@duke.edu
S. Ferrari is with Faculty of Mechanical Engineering at Duke University, Durham, NC, 27708, USA sferrari@duke.edu

tion of a variable x given evidence about the observed variables in the network.

BNs map causal-effect relationships among all relevant variables by learning the joint probability distributions from data and, possibly, heuristic arguments. They can be used for modeling a generic sensor measurement process by considering all of the variables that influence the measurements outcome. These variables that can be classified as follows:

- Sensor Mode (S): The set of sensor parameters chosen to operate the sensor.
- Environment (E): The set of environmental variables influencing sensor measurements.
- Observed Features (F): The set of target characteristics that are obtained from the raw sensor measurements.
- Actual Target Features (T): The set of actual target characteristics that must be inferred from the sensor measurements.

In many applications, such as demining, the measurements obtained from multiple and heterogeneous sensors must be obtained and fused in order to achieve satisfactory classification performance. The BN architecture shown in Fig. 1 [4] can be used to model each sensor using the procedure in [4]. In some cases, the likelihood $P[F \mid T]$ may be given in terms of a known probability density function (PDF), such as the Gaussian distribution used in [2]. But, in general, $P[F \mid T]$ may be unknown and non-Gaussian. Then the BN approach presented in this paper can be used to learn the PDF from available sensor data.

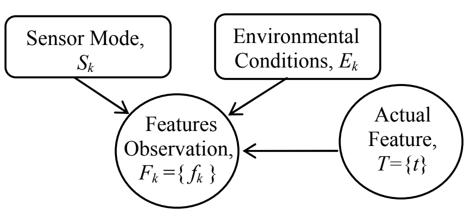

Fig. 1. BN architecture for modeling k^{th} sensor (for $1 \le k \le n$) when a target has one feature, $T = \{t\}$.

B. Sensor Fusion

When two sensors, a and b, are applied to collect measurements from the same feature t, the respective BN models are used to obtain the probability distribution of t over all of its possible states, e.g., $P_1[t = t_i]$ and $P_2[t = t_j]$, where P_k is obtained from the k^{th} sensor, l is the size of the domain of t, and $j = 1, \ldots, l$. The distributions are combined by the D-S rule of evidence combination [9, 10] to produce the fused probability of each state t_k, as follows

$$P_{\oplus}[t_k] = \frac{P_1[t_k]P_2[t_k]}{1 - \sum_{i,j=1\ \&\ i \neq j}^{l} P_1[t_i]P_2[t_j]} \quad (1)$$

For multiple sensors fusion is implemented as follows: first the probabilities obtained by two sensor measurements, P_1 and P_2, are fused to obtain $P_{\oplus}$. Then, $P_{\oplus}$ is fused with the next distribution, P_3, and the iterative process continues until every distribution P_k, $k = 1, \ldots, n$, is incorporated by eq. (1).

C. Review of Information-Theoretic Functions

The discrimination or *cross-entropy* is a measure of "distance" between two probability distributions [11]. Consider the problem of estimating feature t with mutually exclusive and countable l states. Let $P[t]$ and $Q[t]$ be two distributions over the domain of t, or $\{t_i \mid i = 1, \ldots, l\}$, as obtained, for instance, by two different information sources. The cross entropy, D, may be used to determine which measurements can improve the distribution, $Q[t]$, obtained from cursory sensor measurements. The cross entropy of P with respect to Q is,

$$D(P;Q) = \sum_t P[t]\log_2(P[t]/Q[t]), \quad (2)$$

and is lower the closer $P[t]$ is to $Q(t)$.

This measure is always non-negative and although it often is interpreted as the "distance" between P and Q, it is not a proper metric. For example, $D(P;Q) \neq D(Q;P)$, it is not additive, and does not obey the triangle inequality.

The entropy of a discrete random variable is a measure of the uncertainty associated with a random variable, as reflected in its probability distribution, $P(t)$. The entropy of a feature t is defined as,

$$H(t) = -\sum_t P[t]\log_2(P[t]) \quad (3)$$

and it can be shown that conditioning, as due to evidence e from another variable, always reduces entropy, i.e., $H(t \mid e) \le H(t)$ [11].

III. Methodology

A. Problem Formulation

Consider the general classification problem of detecting targets confined to discrete cells indexed by $c = 1, \ldots, c_f$, where c_f is the total number of cells. The state x of each cell, c, has l-possible values and is determined from a set of m features of the same target, $T = \{t^j \mid j = 1, .., m\}$. Let n denote the number of different sensors used to measure the m features of the target cells, producing a set of observation outcomes $F_k = \{f_k^j \mid j = 1, .., m\}$, $k = 1, \ldots, n$, for each feature t^j. Then, the classification problem consists of using measured target features to infer the type of target in the cell, i.e., x. A problem in sensor management is to direct n sensors to a subset of all available cells, and to select the sensor modes, such that the fused classification accuracy can be maximized by a fixed amount of measurements.

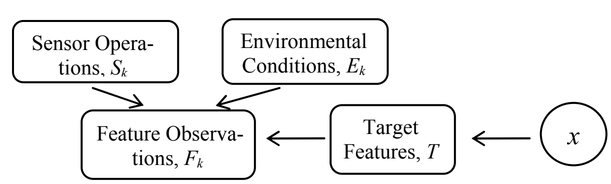

Fig. 2. BN architecture for modeling sensor k (total n sensors) in target classification case

When $m = 1$, the classification problem is reduced to feature inference, as shown in Fig. 1, where the classification goal is to infer the only target feature t. When $m \geq 2$, and the features depend on the target typology, as shown in Fig. 2, the problem is referred to as target classification.

B. Virtual BN Formulation and Problem Solution

The initial BN architecture in Fig. 1 is used to construct a BN model for each sensor type. The virtual BN representing the entire sensor system is comprised of n BN sensor models and of one BN classifier, as shown in Fig. 3. For the feature inference case, it can be assumed that there is only one feature in T, and the node x is eliminated.

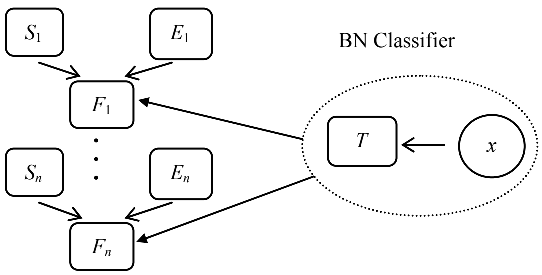

Fig. 3. A virtual BN including n sensor measurements

The virtual BN illustrates the conditional independence relationships between the n sensor measurements F_k, $k = 1, \ldots, n$. As shown in Fig. 3, the n sensor measurements F_k, $k = 1, \ldots, n$, are d-separated by the target feature set T, and, therefore, they are conditionally independent given T. Another advantage of this virtual BN is that it contains all of the conditional probability distributions needed to compute the information objective functions, as shown in the following section.

C. Information Theoretic Objective Functions for Sensor Planning

The expected discrimination gain and expected entropy reduction, are used to assess the expected value of sensor measurements prior to sensor deployment. Let $Z_{k-1} = \{F_1, \ldots, F_{k-1}\}$ and suppose $(k-1)$ different sensors have been used to obtain measurements from each cell. Then, the sensor manager must select the *next* cell that maximizes the information objective function for each possible observation outcome that could be obtained by the k^{th} sensor once it is deployed. Theoretic solutions for feature inference (Fig. 1) and target classification (Fig. 2) are stated below.

1. Feature Inference

In the feature inference case, shown in Fig. 1, the target in cell c can be assumed to have only one feature, $T = \{t\}$, and $F_k = \{ f_k \}$, $k = 1, \ldots, n$. The prior $P_c[t]$ and the conditional probabilities $P_c[f_k \mid t]$ are all known from the BN models. Since the following derivation holds for any cell, the cell subscript c is omitted for brevity, thus $P_c = P$.

The probability of observing Z_k is written in terms of the conditional probability as,

$$P[Z_k] = \sum_T P[Z_k \mid T]P[T] \tag{4}$$

with the summation representing marginalization over T. Since the variables $\{F_k\}_{k=1,\ldots,n}$ are d-separated or conditionally independent given T it follows that,

$$P[Z_k \mid T] = \prod_{i=1}^{k} P[F_i \mid T] \tag{5}$$

and the cell's feature state probabilities are given by,

$$P[T \mid F_i] = \frac{P[T]P[F_i \mid T]}{\sum_U P[U]P[F_i \mid U]} \tag{6-a}$$

$$P[T \mid Z_k] = \frac{P[T]\prod_{i=1}^{k} P[F_i \mid T]}{\sum_U P[U]\prod_{i=1}^{k} P[F_i \mid U]} \tag{6-b}$$

where $U = \{t\}$ is the feature set.

To compute the discrimination gain, the probability distribution for the k^{th} observation after $(k-1)$ observations is,

$$\begin{aligned} P[F_k \mid Z_{k-1}] &= \sum_T P[F_k \mid T, Z_{k-1}]P[T \mid Z_{k-1}] \\ &= \sum_T P[F_k \mid T]P[T \mid Z_{k-1}], \end{aligned} \tag{7}$$

because the variables $\{F_k\}_{k=1,\ldots,n}$ are conditionally independent given the target feature T.

After k observations, the discrimination of the inferred feature distribution with respect to the prior $P[T]$ is computed in terms of $P[T \mid Z_k]$:

$$\begin{aligned} D[Z_k] &= D(P[T \mid Z_k]; P[T]) \\ &= \sum_T P[T \mid Z_k]\log_2(P[T \mid Z_k]/P[T]) \end{aligned}. \tag{8}$$

After k observations, the entropy of the inferred feature is,

$$H(T \mid Z_k) = -\sum_T P[T \mid Z_k]\log_2 P[T \mid Z_k] \tag{9}$$

while after $(k-1)^{\text{th}}$ observations, and before the k^{th} observation is obtained, the expected discrimination gain is,

$$E[D \mid Z_k] = \sum_{F_k} D[Z_k]P[F_k \mid Z_{k-1}] \tag{10}$$

and the expected entropy is:

$$E[H \mid Z_k] = \sum_{F_k} H(T \mid Z_k)P[F_k \mid Z_{k-1}]. \tag{11}$$

It follows that for an individual cell c, the expected discrimination gain is

$$\Delta D(Z_k) = E[D \mid Z_k] - D[Z_{k-1}] \tag{12}$$

and the expected entropy reduction is

$$\Delta H(Z_k) = H(T \mid Z_{k-1}) - E[H \mid Z_k]. \tag{13}$$

$\Delta H(Z_k)$ is always non-negative, while $\Delta D(Z_k)$ can be negative. An important advantage of this approach is that all of the probabilities required, such as $P[T \mid Z_k]$ and $P[F_k \mid Z_{k-1}]$, can be obtained from the virtual BN model in Fig. 3.

2. Target Classification

In the target classification case, T and F_k, with $k = 1, \ldots, n$, are sets containing multiple variables. Thus, $P[F_k \mid T]$ and $P[T \mid x]$ are joint conditional probability distributions. These two conditional probabilities as well as the prior $P[x]$ are all known from the BN model for cell c.

The probability of $P[T]$ is computed as,

$$P[T] = \sum_x P[T \mid x]P[x] \tag{14}$$

and the notation is modified such that T and F_k are marginalized. With this modification, eq. (4) is still applicable to obtain $P[Z_k]$. Then, since the connections of x, T and F_k are serial, x and F_k are d-separated given T. Hence, the following equalities hold,

$$P[F_k \mid x] = \sum_T P[F_k \mid T, x]P[T \mid x] = \sum_T P[F_k \mid T]P[T \mid x] \tag{15}$$

$$P[Z_k \mid x] = \sum_T P[Z_k \mid T, x]P[T \mid x] = \sum_T P[Z_k \mid T]P[T \mid x] \tag{16}$$

where $P[Z_k \mid T]$ can be computed from eq. (5). The probability distribution for the cell typology (or target classification) is given by:

$$\begin{aligned} P[x \mid Z_k] &= \sum_T P[x \mid T]P[T \mid Z_k] \\ &= \frac{P[x]\sum_T P[Z_k \mid T]P[T \mid x]}{\sum_T P[Z_k \mid T]P[T]}. \end{aligned} \tag{17}$$

Equation (7) also holds under the change in the notation of marginalization over all members of T.

After k observations, the discrimination of the target classification variable's distribution with respect to the prior $P[x]$ is computed using $P[x \mid Z_k]$ in (17),

$$\begin{aligned} D[Z_k] &= D(P[x \mid Z_k]; P[x]) \\ &= \sum_x P[x \mid Z_k]\log_2(P[x \mid Z_k]/P[x]) \end{aligned} \tag{18}$$

and the entropy of the classification variable is

$$H(x \mid Z_k) = -\sum_x P[x \mid Z_k]\log_2 P[x \mid Z_k] \tag{19}$$

Equation (10) can be used to compute the expected discrimination, provided $D[Z_k]$ is obtained from eq. (18). Then, the expected entropy after $(k-1)^{\text{th}}$ observations, and before the k^{th} observation is obtained, is

$$E[H \mid Z_k] = \sum_{F_k} H(x \mid Z_k)P[F_k \mid Z_{k-1}] \tag{20}$$

Thus, given the above results, eq. (13) can be used to compute the expected discrimination gain. Also, the expected entropy reduction can now be computed as:

$$\Delta H(Z_k) = H(x \mid Z_{k-1}) - E[H \mid Z_k] \tag{21}$$

The formulas derived in this section allow to compute the expected discrimination gain and expected entropy reduction for target classification using probabilities, such as $P[x \mid Z_k]$ and $P[F_k \mid Z_{k-1}]$, that can all be obtained from the virtual BN model in Fig. 3, similarly to the feature inference case.

D. D-S Sensor Fusion

The previous section provides a framework for combining $n \geq 2$ sensor measurements of the same features. Another approach that is based on the D-S fusion rule is shown in Fig. 4. Here, for any feature t^j, $j = 1, \ldots, m$, the posterior distribution $U_k^j = P[t^j \mid F_k]$ is obtained from (6-a) or from the k^{th} sensor BN model, $k = 1, \ldots, n$. When $k \geq 2$, a total of k posterior distributions U_i^j, $i = 1, \ldots, k$, are needed in order for the D-S fusion method to be applicable, and to produce the distribution $U_\oplus^j = P[t_\oplus^j \mid Z_k]$ as described in Section II.*B*. When the feature number is $m \geq 2$, the fused distribution $U_\oplus^j$ is used as soft evidence for the BN classifier model to compute two distributions, namely, the posterior distribution of target classification variable, $P[x \mid Z_k]$, and the joint probability $P[T_\oplus \mid U_\oplus^1, \ldots, U_\oplus^m]$. Where, the fused feature set is defined as $T_\oplus \equiv \{t_\oplus^j \mid j = 1, \ldots, m\}$. Since $U_\oplus^j = P[t_\oplus^j \mid Z_k]$, for any $j = 1, \ldots, m$, and $T_\oplus = \{t_\oplus^1, \ldots, t_\oplus^m\}$, then $P[T_\oplus \mid U_\oplus^1, \ldots, U_\oplus^m]$ is equal to $P[T_\oplus \mid Z_k]$ or, simply, $P[T \mid Z_k]$, for any $k = 2, \ldots, n$.

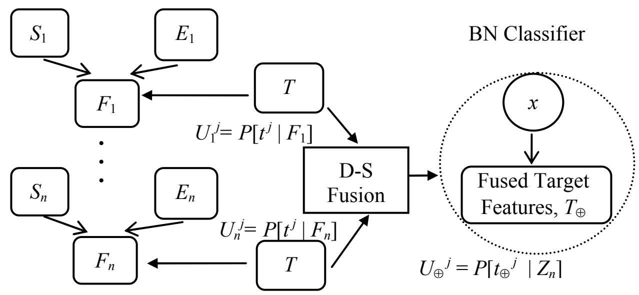

Fig. 4. Feature D-S fusion in classification

In target classification based on sensor fusion, $P[F_k \mid Z_{k-1}]$ can be computed as,

$$\begin{aligned} P[F_k \mid Z_{k-1}] &= \sum_T P[F_k \mid T]P[T \mid Z_{k-1}] \\ &= \sum_T P[F_k \mid T]P[T_\oplus \mid Z_{k-1}]. \end{aligned} \tag{22}$$

The expected discrimination gain and expected entropy reduction can be obtained from eqs. (12) and (21), respectively, provided that $D[Z_k]$ is obtained from eq. (18) and $P[F_k \mid Z_{k-1}]$ is obtained from eq. (22).

The main difference between these two approaches lies in the computation of $P[T \mid Z_k]$. In the non-fusion case, $P[T \mid Z_k]$ is computed from eq. (6-b). In the sensor-fusion case, $P[T \mid Z_k]$ is approximated by combining the D-S fusion rule with the BN classifier model.

IV. Application to Landmine Detection and Classification

The information functions are applied to a demining system comprised of airborne infrared (IR) sensors, ground penetrating radar (GPR) and electromagnetic induction (EMI) sensors mounted on ground vehicles that search for potential targets. The features measured by these sensors are shape, size, depth, and metal content. The targets buried underground must ultimately be classified as either landmines or clutter. Environmental conditions that are heterogeneously distributed over the field influence the individual performance of each sensor at a given location in the workspace. The sensor-planning problem considered in this paper consists of determining an optimal policy that decides optimal search cell sequence by the IR and GPR sensors, given a maximum allowable number of measurements.

A minefield shown in Fig. 5 is generated by the simulation of landmine detection systems developed in [4]. The targets include anti-tank mines, anti-personnel mines, unexploded ordnance, and clutter objects that have been reproduced based on the Ordata Database [12]. Prior information is obtained from the 98 target cells in the minefield in Fig. 5 by a remote (e.g., airborne) IR sensor. The goal is to optimally direct GPR sensors to obtain additional measurements that complement the existing IR measurements and maximize the improvement of fused IR-GPR classification accuracy. It is assumed that the GPR sensor is only allowed to make a fixed number of measurements due to energy and time limitations.

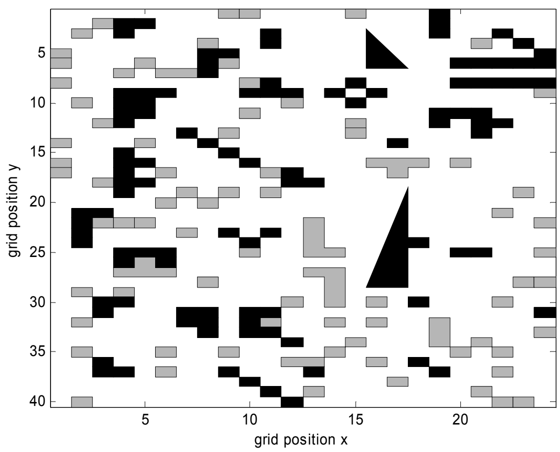

Fig. 5. Prior IR sensor measurements of the minefield obtained by an airplane flying over the region (targets are grey and obstacles are black).

For comparison, the target cells are searched using the following three methodologies:

- Directed Search (DS): advance through the cells in the same order for every frame, taking one measurement over each cell.
- Discrimination Gain Based Search (DGBS): direct the sensor to search the cells with the highest expected discrimination gain, taking one measurement over each cell.
- Entropy Reduction Based Search (ERBS): direct the sensor to search the cells with the highest expected entropy reduction, taking one measurement over each cell.

An IR sensor BN model (Fig. 6), a GPR sensor BN model and a BN classifier are developed and learned from training data, as shown in [4]. These BN models are combined with D-S fusion rule to compute expected discrimination gain and expected entropy reduction.

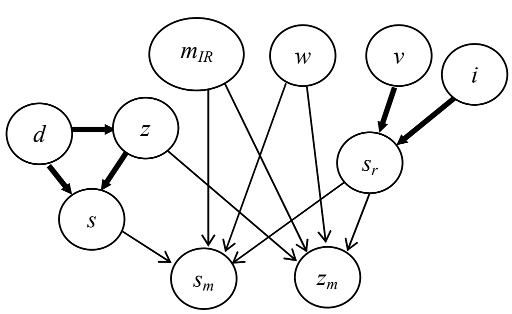

Fig. 6. IR BN sensor model taken from [4], where d is the depth (cm), z the size (cm), s the shape, m the measurement, m_{IR} is the IR mode, s_r:is the soil moisture (%), w is weather, v is vegetation, and i is illumination.

Based on these objective functions the GPR sensor is directed to obtain measurements from a selected cell sequence according to DGBS or ERBS. Figures 7-9 show comparisons of average classification accuracy using the three searching techniques with a fixed number of GPR measurements (shown on the abscissa).

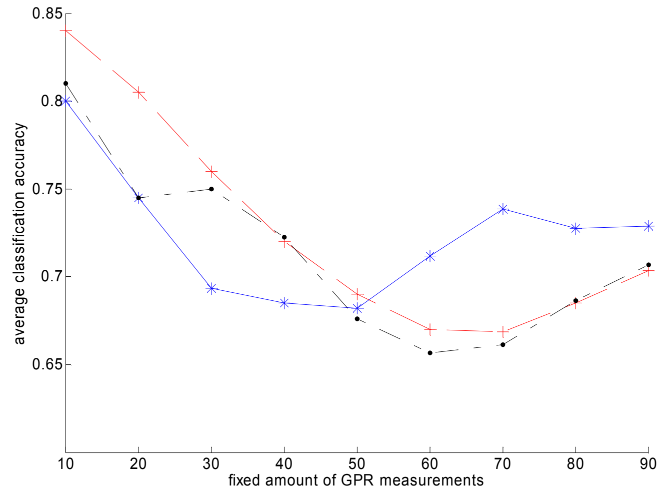

Fig. 7. Average classification accuracy using three searching technique, obtained by averaging 10-trials.

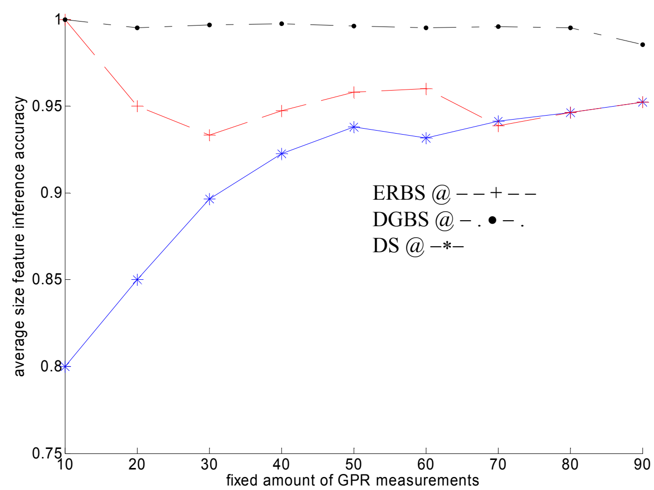

Fig. 8. Average size (z) feature-inference accuracy using three searching techniques, obtained by averaging 10-trials.

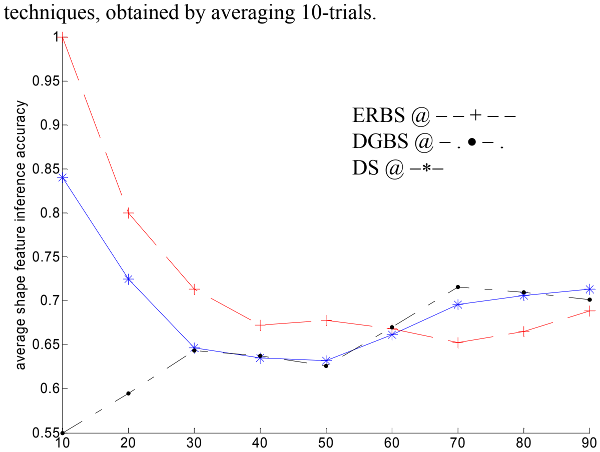

Fig. 9. Average shape (s) feature-inference accuracy using three searching techniques, obtained by averaging 10-trials.

As can be seen from Fig. 7, when up to 50 GPR measurements are taken, both a high expected discrimination gain and a high expected entropy reduction lead to high target classification accuracy. With less than 40 cell measurements, ERBS obtains better classification accuracy than DGBS. In both ERBS and DGBS, the classification accuracy unexpectedly increases after 60 fixed GPR measurements. This increase may be caused by one of the following factors: (1) random noise in sensor measurements, (2) BN modeling errors due to noisy sensor training data, or (3) approximation in computing objective functions due to the D-S rule.

The two information-theoretic objective functions are also implemented in feature inference examples. In Fig. 8, both DGBS and ERBS achieve high accuracy for the inference of the size-feature of buried targets. Although in this case DGBS is slightly better than ERBS, it should be noted that the classification accuracy difference is very small, and a better size inference does not necessarily result in a better overall target classification, since the confidence level of inferred size posterior is not yet considered. Figure 9 shows that the expected entropy-reduction objective function is much better than the expected discrimination-gain objective function for the inference of the shape feature, with less than 30 GPR measurements. The fact that average shape inference accuracy unexpectedly increases in DGBS shows that expected discrimination-gain objective function does not work for shape inference.

Based on the results shown in Fig. 7-9, it is concluded that expected entropy reduction is generally preferable to the expected discrimination gain for the purpose of sensor planning, particularly in the case of target classification.

V. Conclusions and Future Work

Two information theoretic functions, namely, expected discrimination gain and expected entropy reduction, are demonstrated and compared for feature inference and target classification. The equations to compute these objective functions are derived in terms of posterior and prior probability distributions that are available from BN models of the sensor measurement processes. Another significant contribution is that the D-S fusion rule is embedded in these computations, thereby accounting for the expected value of the complementarity of measurements obtained from multiple and heterogeneous sensors. A virtual BN framework is proposed to integrate different sensor models and classifiers, as well as sensor fusion.

A simulated demining system is used to compare search methods based on the two objective functions with direct search. The results suggest that both objective functions are efficient for sensor planning aimed at optimizing classification performance. In these simulations, expected entropy reduction is found to be particularly effective and reliable when sensor fusion is performed for the purpose of target classification. On-going work by the authors is considering planning the actions of both the sensors and their platforms to maximize the expected value of information, and minimize the cost associated with the use of sensor and platform resources. Another direction that is being investigated is the use of discrimination gain or entropy reduction in sensor management for properly detecting and classifying dynamic targets.

Acknowledgment

This work was supported by the Office of Naval Research Young Investigator Program (Code 321) and, in part, by the National Science Foundation under CAREER award ECS 0448906.

References

[1] W. Schmaedeke, "Information based sensor management," in *Proceedings of SPIE, Signal processing, sensor fusion, and target recognition* II, pages 156-- 164. SPIE, April 1993.

[2] K. Kastella, "Discrimination gain to optimize detection and classification," *IEEE Trans. Systems, Man, and Cybernetics—Part A:Systems and Humans*, vol. 27, no. 1, pp. 112-116, Jan. 1997.

[3] F. Zhao, J. Shin and J. Reich, "Information-driven dynamic sensor collaboration," *IEEE Signal Processing Mag*., 19 (2002) pp. 61 --72.

[4] S. Ferrari and A. Vaghi, "Demining Sensor Modeling and Feature-Level Fusion by Bayesian Networks," *IEEE Sensors Journal*, Vol. 6, No. 2, pp. 471-483, April 2006,

[5] D. M. Buede and P. Girardi, "A target identification comparison of Bayesian and Dempster-Shafer multisensor fusion," *IEEE Trans. Syst., Man, Cybern.—Part A:Systems and Humans*, vol. 27, no. 5, pp. 569–577, Sept. 1997.

[6] H. Leung and J.Wu, "Bayesian and Dempster-Shafer target identification for radar surveillance," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 36, no. 4, pp. 432–447, Apr. 2000.

[7] J. K. Paik et al., "Image processing-based mine detection techniques using multiple sensors: A review," *Subsurf. Sens. Technol. Appl., Int. J.*, vol. 3, no. 3, pp. 203–252, 2002.

[8] F. V. Jensen, *Bayesian Networks and Decision Graphs*. Springer. 2001, ch. 1.

[9] A. P. Dempster, N. M. Laird, and D. B. Rubin, "A generalization of Bayesian inference," *J. Roy. Stat. Soc. B*, vol. 39, pp. 1–39, 1968.

[10] G. Shafer, *A Mathematical Theory of Evidence*. Princeton, NJ: Princeton Univ. Press, 1976.

[11] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley and Sons, Inc., New York, N.Y., 1991.

[12] Naval Explosive Ordnance Disposal Technology Division, Ordata Online, U.S. DoD Humanitarian Demining Program, 2004, (http://maic.jmu.edu/ordata/mission.asp).