Waveform design by Task-Specific Information

Hyoung-soo Kim
Department of Electrical and Computer Engineering
University of Arizona
Tucson, Arizona 85721
Telephone: (520) 621–4025
Email: hskim@ece.arizona.edu

Nathan A. Goodman
Department of Electrical and Computer Engineering
University of Arizona
Tucson, Arizona 85721
Telephone: (520) 621–4462
Email: goodman@ece.arizona.edu

***Abstract*— We propose a new waveform design metric called Task-Specific Information (TSI). As its name suggests, TSI is an information-theoretic metric formulated for a specific task or set of tasks. For example, in this paper we consider the problem of correctly classifying a linear system from a set of known alternatives. The TSI metric directly quantifies information about the hypotheses while previous information-based approaches quantify mutual information related to potential system realizations. Thus the TSI framework is a more direct measure of task-related performance. In addition, for the problem statement considered in this paper it is possible to find a globally optimum solution through gradient-based search techniques. We compare the performance of TSI and conventional waveform designs.**

I. INTRODUCTION

Recently, the problem of radar waveform design for improved detection, parameter estimation, and/or classification has been an intensively studied topic. One result that has been explored recently is the one by Bell, who developed information-theoretic principles where the waveform optimization is achieved by maximizing the mutual information between a Gaussian ensemble of extended targets and the received signal [1]. This approach yields a spectral waterfilling (WF) strategy for obtaining the optimum waveform's energy spectrum. In[2], the approach in [1] was applied to the multiple target discrimination problem by defining a probability-weighted spectral variance across the transfer functions of the various target hypotheses and using this spectral variance in the waterfilling operation. To the extent that accurately estimating the target transfer function yields good detection performance, this approach is valid. However, we are motivated to find a more effective and direct information measure optimized for the specific task of target discrimination. This metric should be directly related to target class identification, not necessarily to estimation of the target spectrum.

The task-specific information (TSI) concept was first proposed as an analysis tool in [3][4] where it was used to optimize features that would be extracted for the final task of recognition in imaging systems. We now develop and apply TSI for waveform design with the goal of discriminating between known target hypotheses. We introduce a virtual source variable $\mathbf{x}$, which is in the form of an indicator vector that selects the true target transfer function from among the ensemble of target hypotheses. The TSI formulation maximizes the mutual information between the received measurements and the source variable $\mathbf{x}$ rather than the measurements and the transfer functions. If the frequency-domain vector of observations is $\mathbf{y}$, then TSI is defined by $I(\mathbf{y};\mathbf{x}) = H(\mathbf{x}) - H(\mathbf{x}|\mathbf{y})$ where $H(\cdot)$ is entropy. For example, if the goal is to correctly identify the target from among four equally likely hypotheses, then the maximum task-specific information that can be extracted is $H(\mathbf{x}) = 2$ bits, which is consistent with the four-class recognition problem.

We show two different simulation results to compare the performance of TSI-based waveform design with other waveform design methods. The first result is generated within the context of a cognitive radar system with the ability to update waveform transmissions based on prior measurements. Thus, a cognitive radar operates in a closed-loop manner. The performance metric for the cognitive radar system is the average number of illuminations required to make a decision with a prescribed probability of error. The second result is simply the probability of error achieved by a system performing maximum likelihood detection based on measurements from a single transmission.

The paper is organized as follows. In the next section, the structure of our system measurement model is described along with the definition of TSI and preliminaries on sequential hypothesis testing [5][6]. In Section 3, we show the gradient of TSI with respect to waveform frequency-domain coefficients in an energy-constrained optimization. In Section 4, the performances of the proposed TSI-based waveform and other waveforms are compared through computer simulation results. Conclusions are made in Section 5.

II. SYSTEM MODEL

Consider a discrete, frequency-domain measurement model where the l^{th} measurement is

$$y(l) = w(l)g(l) + n(l). \qquad (1)$$

In (1), $w(l)$ is the l^{th} frequency-domain waveform coefficient, $g(l)$ is the l^{th} target transfer function coefficient, and $n(l)$ is additive Gaussian noise. For measurements taken at $\mathcal{L}$ frequencies, the $\mathcal{L}\times 1$ measurement vector is then $\mathbf{y} = \mathbf{W}\mathbf{g}+\mathbf{n}$ where $\mathbf{y} = [y(1), y(2), \ldots, y(\mathcal{L})]^T$, $\mathbf{g} = [g(1),\ g(2),\ \ldots,\ g(\mathcal{L})]^T$, the $\mathcal{L}\times\mathcal{L}$ waveform matrix $\mathbf{W}$ is diagonal with entries $w(1)$, $w(2)$, $\ldots w(\mathcal{L})$, and $\mathbf{n}$ is a vector of independent Gaussian noise samples. For the case where we are trying to identify one of

978-1-4244-5812-7/10/$26.00

$\mathcal{M}$ known hypotheses, the target transfer function can take one of $\mathcal{M}$ possible realizations. Let these known realizations be $\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_{\mathcal{M}}$, and define the $\mathcal{L} \times \mathcal{M}$ target transfer matrix as $\mathbf{G} = [\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_{\mathcal{M}}]$. The measurements can now be described by $\mathbf{y} = \mathbf{WGx} + \mathbf{n}$. The vector $\mathbf{x}$ is an indicator vector $\{\mathbf{x} = \mathbf{e}_m | m = 1, 2, \ldots, \mathcal{M}\}$ where $\mathbf{e}_m$ is a vector with a one in the m^{th} position and zeros elsewhere. For example, if the first hypothesis is the true hypothesis, then $\mathbf{x} = \mathbf{e}_1$, which selects the first column of $\mathbf{G}$. The pdf of the measurement vector is

$$\begin{aligned} & p(\mathbf{y}|\mathbf{x}) \\ = \quad & \frac{1}{(2\pi\sigma^2)^{\frac{L}{2}}} \exp\left[-\frac{(\mathbf{y} - \mathbf{WGx})^T(\mathbf{y} - \mathbf{WGx})}{2\sigma^2}\right] \\ = \quad & \frac{1}{(2\pi\sigma^2)^{\frac{L}{2}}} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{\mathcal{L}}(y(i) - w(i)\{\mathbf{Gx}\}(i))^2\right]. \end{aligned}$$

There are $\mathcal{M}$ possible realizations of $\mathbf{x}$, one realization for each target hypothesis. Thus, the virtual source variable $\mathbf{x}$ encodes the discrimination task into the signal model. If the hypotheses are equally likely, then the task-specific entropy prior to taking any measurements is $H(\mathbf{x}) = log_2(\mathcal{M})$ bits. While other information-based waveform design methods are based on the mutual information between $\mathbf{y}$ and the target transfer function ensemble $\mathbf{G}$, the introduction of a source variable admits a new task-specific metric.

Using the above measurement model, TSI is defined as $TSI = I(\mathbf{y}; \mathbf{x})$, and the optimization problem is now

$$\begin{aligned} \max \quad & TSI = I(\mathbf{y}; \mathbf{x}) \\ \text{s.t.} \quad & Tr[\mathbf{WW}^T] \le E_w, \end{aligned} \tag{2}$$

where $Tr[\cdot]$ is the matrix trace and E_w is a constraint on the allowable waveform energy.

A. Sequential Hypothesis Testing for Cognitive Radar

In [2], the authors propose an iterative feedback target classification method integrated with sequential hypothesis testing. The iterative method combined with the ability to adapt subsequent transmissions is based on the concept of cognitive radar, which was first described in [7]. In the application considered in this paper, the radar's understanding of the scenario is quantified by knowledge of the target transfer functions in the matrix $\mathbf{G}$ as well as the probability P_i that the i^{th} hypothesis is true. Based on this knowledge, a customized waveform is transmitted and measurements are collected. Based on these measurements, sequential hypothesis testing allows for one of $\mathcal{M}+1$ possible courses of action: we can select one of the $\mathcal{M}$ hypotheses or we can elect to make another transmission. If another transmission is made, then the probabilities P_i are first updated using Bayes' Rule, then the updated probabilities are used to enhance the waveform design incorporating what was learned on the prior transmission. Define the likelihood ratio between the i^{th} and j^{th} hypotheses after the k^{th} transmission/reception as

$$\lambda_{i,j}^k = \frac{\prod_{m=1,\ldots,k} p(\mathbf{y}_m|\mathbf{x} = \mathbf{e}_i) \cdot P_i}{\prod_{l=1,\ldots,k} p(\mathbf{y}_l|\mathbf{x} = \mathbf{e}_j) \cdot P_j}. \tag{3}$$

In (3), $\mathbf{y}_m$ is the measurement vector obtained from the mth transmission, P_i is the prior probability of the i^{th} hypothesis prior to taking the first measurement, and the pdf's for each measurement are multiplied by virtue of the independence between noise samples taken on different transmissions. If after the k^{th} transmission, the likelihood ratio $\lambda_{i,j}^k$ is greater than $\frac{1-P_e}{P_e}$ for all j not equal to i, then we decide in favor of i^{th} hypothesis.

III. Waveform Design with Task-Specific Information

In the optimization problem of (2), TSI is used as the objective function. Let TSI_k be the TSI obtained due to the k^{th} transmission. We use a simple gradient search method to find the waveform energy spectrum that maximizes TSI at the k^{th} transmission. The derivative of TSI_k with respect to the i^{th} frequency component of k^{th} waveform, $w_k(i)$ is [4]

$$\begin{aligned} \frac{\partial TSI_k}{\partial w_k(i)} &= \frac{\partial I(\mathbf{y}_{(1:k-1)}; \mathbf{x})}{\partial w_k(i)} + \frac{\partial I(\mathbf{y}_{(1:k)}; \mathbf{x}|\mathbf{y}_{(1:k-1)})}{\partial w_k(i)} \\ &= \frac{\partial I(\mathbf{y}_{(1:k)}; \mathbf{x}|\mathbf{y}_{(1:k-1)})}{\partial w_k(i)} \\ &= \frac{\partial TSI_{k|k-1}}{\partial w_k(i)} \end{aligned}$$

since the past TSI information $I(\mathbf{y}_{(1:k-1)}; \mathbf{x})$ does not depend on the current waveform coefficient $w_k(i)$. After differentiating $TSI_{k|k-1}$ with respect to $w_k(i)$ and doing some algebra, we obtain

$$\begin{aligned} \frac{\partial TSI_{k|k-1}}{\partial \widetilde{\mathbf{W}}_k} = \frac{\mathbf{W}_k}{\sigma^2} E_{\mathbf{x},\mathbf{y}_k}\Big[p(\mathbf{y}_{(1:k-1)}|\mathbf{x})\Big[diag\{\mathbf{Gx}\} \cdot \mathbf{Gx} \\ - E_{\mathbf{x}|\mathbf{y}_{(1:k)}}\left[diag\{\mathbf{Gx}\}\right] \cdot E_{\mathbf{x}|\mathbf{y}_{(1:k)}}\left[\mathbf{Gx}\right]\Big]\Big] \end{aligned} \tag{4}$$

where

$$\frac{\partial TSI_{k|k-1}}{\partial \widetilde{\mathbf{W}}_k} = \begin{bmatrix} \frac{\partial TSI_{k|k-1}}{\partial w_k(1)} \\ \frac{\partial TSI_{k|k-1}}{\partial w_k(2)} \\ \vdots \\ \frac{\partial TSI_{k|k-1}}{\partial w_k(\mathcal{L})} \end{bmatrix}.$$

For numerical programming, (4) can be written as

$$\begin{aligned} & \frac{\partial TSI_{k|k-1}}{\partial \widetilde{\mathbf{W}}_k} \\ & = \frac{\mathbf{W}_k}{N_{total} \cdot \sigma^2} \sum_{\mathbf{y}_0, \mathbf{x}_0} \Big[p(\mathbf{y}_{(1:k-1)}|\mathbf{x}_0)\left\{diag\{\mathbf{Gx}_0\} \cdot \mathbf{Gx}_0\right. \\ & \quad \left. - E_{\mathbf{x}|\mathbf{y}_{(1:k-1)},\mathbf{y}_0}\left[diag\left\{\mathbf{Gx}\right\}\right] \cdot E_{\mathbf{x}|\mathbf{y}_{(1:k-1)},\mathbf{y}_0}\left[\mathbf{Gx}\right]\Big]\right. \end{aligned} \tag{5}$$

where N_{total} is the number of Monte Carlo realizations necessary for estimating the required expected value,

$$E_{\mathbf{x}|\mathbf{y}_{(1:k-1)},\mathbf{y}_0}[\mathbf{Gx}] = \frac{\sum_{j=1}^{M}\{G\mathbf{e_j}\}p(\mathbf{y}_{(1:k-1)}|\mathbf{x}=\mathbf{e_j})p(\mathbf{y}_0|\mathbf{x}=\mathbf{e_j})P(\mathbf{x}=\mathbf{e_j})}{\sum_{j=1}^{M}p(\mathbf{y}_{(1:k-1)}|\mathbf{x}=\mathbf{e_j})p(\mathbf{y}_0|\mathbf{x}=\mathbf{e_j})P(\mathbf{x}=\mathbf{e_j})},$$

and $diag\{\mathbf{Gx}\}$ is a diagonal matrix including the vector $\mathbf{Gx}$ on the main diagonal. Equation (4) does not include the waveform energy constraint E_w. We next consider an $\mathcal{L}$-dimensional spherical coordinate transformation for defining an unconstrained optimization problem.

A. $\mathcal{L}$-dimensional spherical coordinate

The optimum waveform that meets the energy constraint will lie on an $\mathcal{L}$-dimensional spheroid. We can exploit this fact by converting the waveform to an $\mathcal{L}$-dimensional spherical coordinate system according to

$$\begin{aligned} w_k(1) &= \sqrt{E_w}\cos\theta_k(1) \\ w_k(2) &= \sqrt{E_w}\sin\theta_k(1)\cos\theta_k(2) \\ &\vdots \\ w_k(\mathcal{L}-1) &= \sqrt{E_w}\prod_{i=1}^{\mathcal{L}-2}\sin\theta_k(i)\cos\theta_k(\mathcal{L}-1) \\ w_k(\mathcal{L}) &= \sqrt{E_w}\prod_{i=1}^{\mathcal{L}-1}\sin\theta_k(i) \end{aligned} \quad (6)$$

where $E_w = \sum_{i=1}^{\mathcal{L}}(w_k(i))^2$. In the spherical coordinate system, if we fix the energy constraint E_w, a spherical search space is established defined by $(\mathcal{L}-1)$ angles and the radius $\sqrt{E_w}$. Alternatively, we can view the spherical coordinate strategy as a projection of the original design parameters $[w_k(1),\ w_k(2),\ \ldots, w_k(\mathcal{L})]^T$ onto the spherical surface. The unconstrained optimization problem can be solved by a gradient search algorithm. After obtaining the optimal angles, we convert them to the $\mathcal{L}$-dimensional waveform by (6). For this unconstrained optimization, we need to define the Jacobian matrix

$$\mathbf{J}(\theta_1,\theta_2,\ldots,\theta_{\mathcal{L}-1},E_w) = \begin{pmatrix} \frac{\partial w_k(1)}{\partial\theta_k(1)} & \frac{\partial w_k(1)}{\partial\theta_k(2)} & \cdots & \frac{\partial w_k(1)}{\partial\theta_k(\mathcal{L}-1)} & \frac{\partial w_k(1)}{\partial E_w} \\ \frac{\partial w_k(2)}{\partial\theta_k(1)} & \frac{\partial w_k(2)}{\partial\theta_k(2)} & \cdots & \frac{\partial w_k(2)}{\partial\theta_k(\mathcal{L}-1)} & \frac{\partial w_k(2)}{\partial E_w} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{\partial w_k(\mathcal{L})}{\partial\theta_k(1)} & \frac{\partial w_k(\mathcal{L})}{\partial\theta_k(2)} & \cdots & \frac{\partial w_k(\mathcal{L})}{\partial\theta_k(\mathcal{L}-1)} & \frac{\partial w_k(\mathcal{L})}{\partial E_w} \end{pmatrix}. \quad (7)$$

By applying (6) to (7) with the fixed energy constraint E_w, we get the following lower diagonal Jacobian matrix,

$$\mathbf{J}(\theta_1,\theta_2,\ldots,\theta_{\mathcal{L}-1}) = \begin{pmatrix} \frac{\partial w_k(1)}{\partial\theta_k(1)} & 0 & \cdots & 0 \\ \frac{\partial w_k(2)}{\partial\theta_k(1)} & \frac{\partial w_k(2)}{\partial\theta_k(2)} & 0 & 0 \\ \vdots & \vdots & \ddots & 0 \\ \frac{\partial w_k(\mathcal{L})}{\partial\theta_k(1)} & \frac{\partial w_k(\mathcal{L})}{\partial\theta_k(2)} & \cdots & \frac{\partial w_k(\mathcal{L})}{\partial\theta_k(\mathcal{L}-1)} \end{pmatrix}. \quad (8)$$

We then get $\frac{\partial TSI_{k|k-1}}{\partial\mathbf{\Theta_k}} = \mathbf{J}(\theta_1,\theta_2,\ldots,\theta_{\mathcal{L}-1})^T\cdot\frac{\partial TSI_{k|k-1}}{\partial\widetilde{\mathbf{W}}_k}$ where $\frac{\partial TSI_{k|k-1}}{\partial\mathbf{\Theta_k}} = [\frac{\partial TSI_{k|k-1}}{\partial\Theta_k(1)}, \frac{\partial TSI_{k|k-1}}{\partial\Theta_k(2)}, \ldots, \frac{\partial TSI_{k|k-1}}{\partial\Theta_k(\mathcal{L}-1)}]^T$. Now, we find a maximum mutual information with respect to waveform matrix. The maximum information can be tracked by following gradient update programming.

$$\mathbf{\Theta}_{k,j+1} = \mathbf{\Theta}_{k,j} + \lambda\left[\frac{\partial TSI_{k|k-1}}{\partial\mathbf{\Theta}_k}\right]_j \quad (9)$$

where $\mathbf{\Theta}_{k,j} = [\theta_{k,j}(1),\theta_{k,j}(2),\ldots,\theta_{k,j}(\mathcal{L})]^T$ and $0 \le \theta_{k,j}(l) \le \pi/2$ where k is an illumination index which is used by sequential hypothesis test algorithm and j is an update-index of the gradient search method. This gradient update method is an unconstrained optimization since the energy constraint is already applied in the Jacobian matrix. $\frac{\partial TSI_{k|k-1}}{\partial\mathbf{\Theta}_k}$ is defined only on the energy constrained sphere. $\theta_{k,j}(l)$ is a non-negative angle less than or equal to $\pi/2$ due to the global optimality characteristic of TSI approach. In the next section, we see that this simple gradient search method can find the global optimum of the TSI objective function.

B. Global Optimality

An important feature of the TSI-based waveform design metric is the ability to find a global maximum. Though we will see that there exist many local maxima, all of the maxima have the same TSI value. In the $\mathcal{L}$-dimensional spherical region, we have $2^{\mathcal{L}}$ quadrants and each quadrant has always one local maximum. The waveform parameters $[w_k(1),\ w_k(2),\ \ldots, w_k(\mathcal{L})]^T$ are all positive in the first quadrant, and the other quadrants vary only by sign changes of waveform parameters. Our results indicate that the waveform coefficients' individual signs are not relevant to the optimal solution; only the energy allocated to each frequency coefficient is important. Figure 1 shows an example of TSI evaluated over the entire design space. In Figure 1, the waveform is defined by only three coefficients ($\mathcal{L} = 3$) and there are four equally likely target hypotheses. The maximum TSI possible is two bits, and the waveform design space is a sphere with radius $\sqrt{E_w}$. Thus, the waveform design is defined by two angles that define a point on the sphere. TSI is shown versus those angles in Figure 1. Each quadrant has only one optimum and we have eight symmetric quadrants in Figure 1. In the next section, we search the first quadrant for the optimal waveform and compare the performance with other waveform designs.

IV. Simulation Result

In this section we present simulation results that indicate the potential benefit of using TSI as a waveform design metric. We first evaluate the average number of transmissions required to achieve target classification based on multiple transmissions that are optimized in response to prior measurements. Figure 2 shows the average number of transmissions for the TSI-based waveform, an information-based waveform where the mutual information metric is between the received measurements and the variance of the target transfer functions (labeled as the waterfilling waveform), and a wideband waveform (energy

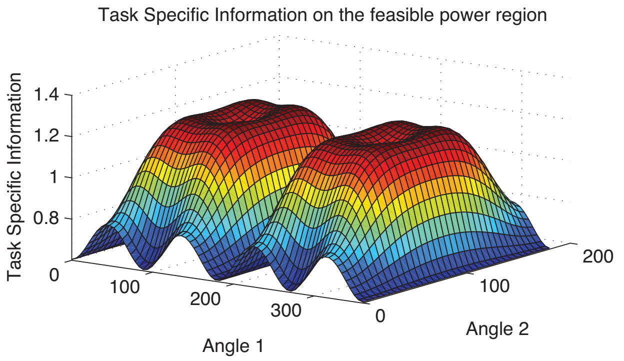

Fig. 1. Plot of TSI on the feasible energy region

allocation is equal across all coefficients). The probability of error used to define the sequential testing thresholds was $P_e = 0.1$, and the number of target hypotheses was $\mathcal{M} = 4$. The waveform dimension was $\mathcal{L} = 10$, the measurement noise energy was $\sigma^2 = 1$, the waveform energy allocation varied from 10^{-2} to 10^1 and 100 deterministic target sets are used for the mean number of illumination. The target transfer functions were randomly generated, but once generated were treated as known for the duration of the classification experiment. Both the TSI-based waveform and the waterfilling waveform outperform the wideband waveform. This fact is easy to understand since these two waveforms are the only ones that take target characteristics into account. The TSI-based waveform requires the fewest number of illuminations among the four waveforms. Figure 2 shows benefit in the introduction of task-specific source variables that are used directly in the optimization metric. However, a more striking performance difference is seen in Figures 3, 4, and 5. Figures 3 and 4 show results for 10-tap and 5-tap waveforms, respectively, with non-colored target sets that are randomly generated. Figure 5 shows the result for 10 tap waveform with colored target sets. The result of colored deterministic target sets shows the wider performance gap in the plot. In this case, each waveform is calculated once, a single transmission is made, and because the hypotheses are equally likely, decisions are based on maximum likelihood detection.

V. Conclusion

A new waveform design method based on the concept of task-specific information is introduced. To evaluate TSI, it is necessary to introduce source variables that directly encode the task at hand. For example, in the $\mathcal{M}$-class identification problem, the source variable is an indicator vector that takes one of $\mathcal{M}$ possible realizations. By optimizing over the mutual information between the measurements and this source variable rather than the mutual information between the measurements and target transfer functions, it is possible to further improve performance beyond what has been demonstrated in earlier work. The TSI-based objective function has a single maxima in each quadrant, and we showed that we need only search over a single quadrant of the design space; thus, gradient-based search techniques lead to a globally optimum solution. The performance improvement of the TSI-based approach is evident in our simulation results.

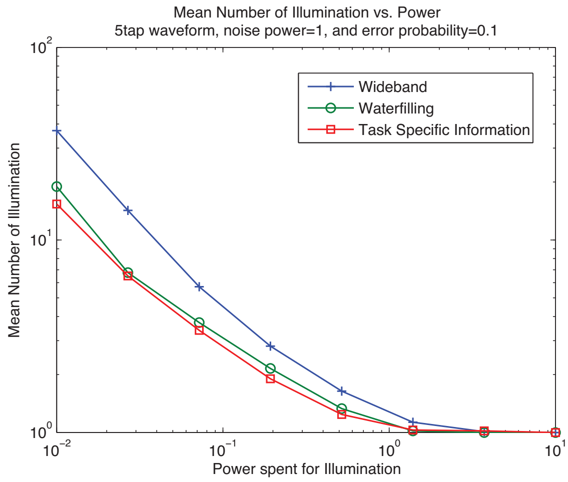

Fig. 2. Performance comparison of various waveform design methods in cognitive radar system in case of 10-tap waveform

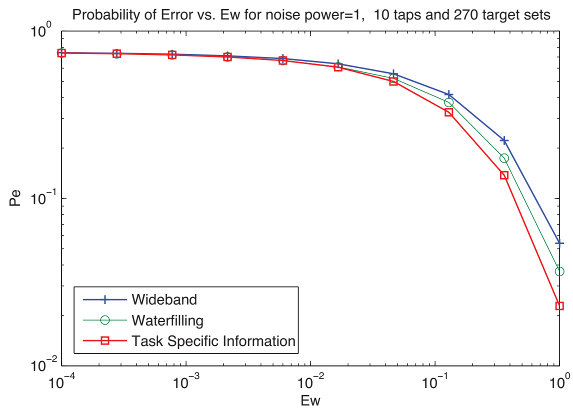

Fig. 3. Performance comparison of various waveform design methods in maximum likelihood detection in case of 10-tap waveform

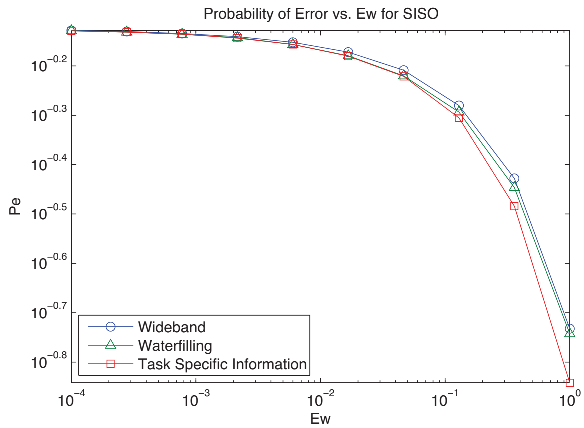

Fig. 4. Performance comparison of various waveform design methods in maximum likelihood detection in case of 5-tap waveform

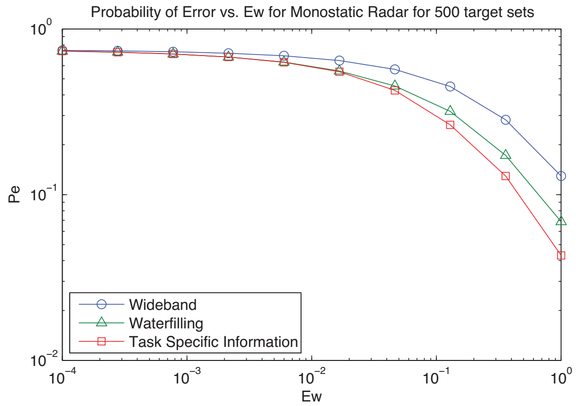

Fig. 5. Performance comparison of various waveform design methods in maximum likelihood detection in case of 10-tap waveform with colored target sets

Acknowledgment

This work was supported in part under AFOSR contract #FA9550-07-1-0182 and ONR contract #N000140910338.‘

References

[1] M.R. Bell, "Information theory and radar waveform design," *Information Theory, IEEE Transactions on*, vol. 39, no. 5, pp. 1578–1597, Sep 1993.

[2] N.A. Goodman, P.R. Venkata, and M.A. Neifeld, "Adaptive waveform design and sequential hypothesis testing for target recognition with active sensors," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 1, no. 1, pp. 105–113, June 2007.

[3] Mark A. Neifeld, Amit Ashok, and Pawan K. Baheti, "Task-specific information for imaging system analysis," *Journal of the Optical Society of America*, vol. 24, no. 12, pp. B25–B164, 2007.

[4] Pawan K. Baheti and Mark A. Neifeld, "Recognition using information-optimal adaptive feature-specific imaging," *Journal of the Optical Society of America*, vol. 26, pp. 1055–1070, 2009.

[5] A. Wald, "Sequential tests of statistical hypotheses," *The Annals of Mathematical Statistics*, vol. 16, no. 2, pp. 117–186, 1945.

[6] P. Armitage, "Sequential analysis with more than two alternative hypotheses, and its relation to discriminant function analysis," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 12, no. 1, pp. 137–144, 1950.

[7] S. Haykin, "Cognitive radar: a way of the future," *Signal Processing Magazine, IEEE*, vol. 23, no. 1, pp. 30–40, Jan. 2006.