Optimal Waveform Selection For Target Classification

Sameh M. Sowelam and Ahmed H. Tewfik
Dept. of Electrical Engineering
University of Minnesota, Minneapolis, MN 55455
e-mail: tewfik@ee.umn.edu

ABSTRACT

This paper studies the design of a set of outgoing radar signals to discriminate between two target classes. We model the reflectivity function of each target by a two-dimensional stochastic process to account for uncertainties and propagation effects. The signals are selected to minimize the expected number of transmissions that are needed to guarantee a given confidence level in the classification decision. We argue that this goal can be achieved by selecting the signals that maximize the *Kullback-Liebler information number* between the two target classes. We illustrate our approach with a particular model. We show that for this model, the optimal set of waveforms can be designed off-line and depends on both the statistics of the reflectivity functions of the targets in both classes and the observation noise level.

1 INTRODUCTION

We consider the problem of identifying moving rigid targets from their wideband radar returns. The target is illuminated with a sequence of pulses and the returns are sequentially processed to arrive the classification decision. Each target is identified by its associated reflectivity function. The reflectivity function $\tilde{D}_c(r,v)$ of an object describes the reflective properties of each elemental point on the target as a function of its range r and radial velocity v. It is more convenient to apply the following transformation to obtain $D_c(x,y)$ from $D_c(r,v)$:

$$x = 2r/c, \qquad y = c - v/c + v \tag{1}$$

where x is the two-way delay to the target at range r, y is a stretch factor and c is the velocity of propagation. The reason for this transformation is as follows. Consider a point target at range r and radial velocity v. Suppose a signal $s(t)$ is transmitted. The received signal after reflection from the object is given by [1]:

$$e(t) = As(y(t-x)) \tag{2}$$

where x and y are given in (1) and A is a constant that depends on the range of the object, its reflectivity properties, frequency of operation among other factors. We will assume that we operate in a high frequency range where the reflectivity remains unchanged over a wide range of frequencies. In narrowband radar the effect of the y factor on the envelope is ignored while its effect on the phase appears as a Doppler frequency shift. In the above equation $D_c(x,y)$ is the reflectivity of a point target located at range $cx/2$ and moving with radial velocity $(1-y)c/(1+y)$ at time $x/2$. The two dimensional function $D_c(x,y)$ then describes a continuum of point targets at different ranges and with different radial velocities. Throughout we will assume that the pulse repetition rate is such that $D_c(x,y)$ remains constant during illumination.

Equation (2) describes the exact broadband return, rather than the usual narrowband approximation. In wideband radar the targets are much larger than the radars resolution cell in range and cross-range. Therefore, one can describe the return $e(t)$ of a transmitted signal $s(t)$ in such a scenario as follows:

$$e(t) = \int_0^\infty dy \int_{-\infty}^\infty dx\, D_c(x,y)s(y(t-x)). \tag{3}$$

In the above equation $D_c(x,y)$ is the reflectivity of a point target located at range $cx/2$ and moving with radial velocity $(1-y)c/(1+y)$ at time $x/2$. The two dimensional function $D_c(x,y)$ then describes a continuum of point targets at different ranges and with different radial velocities. For simplicity, we will assume in the remainder of this paper that the pulse repetition rate is such that $D_c(x,y)$ remains constant during illumination. The general case can be solved using the techniques that we outline here and is described in a forthcoming paper.

We will use (3) to determine the outgoing signals $s(t)$ to efficiently discriminate between targets of different classes. The targets of interest are in general divided into M classes. The reflectivity functions of targets in a particular class obey the same probabilistic description. These reflectivity functions correspond to the class signatures that are used in pattern recognition based target discrimination methods [2].

We set up the problem as a test of statistical hypothesis that is solved sequentially using a sequential test procedure [3]. Our objective then is to construct the outgoing pulse stream to achieve the minimum average number of transmissions for a given probability of error. Previous attempts at sequential radar classification employed a single pulse waveform to obtain multiple observations of the target [6]. We will show that improvement is possible by using different pulses in each transmission depending on the stage of the classification process. We model the reflectivity functions of each class as a two-dimensional stochastic process and determine the signals that provide maximum discrimination between classes.

For purposes of this paper we concentrate on the binary classification problem. Our results extend in a straightforward manner to the multi-class case.

2 Finite Dimensional Approximation

In this section we introduce a discretization of (3) that will considerably simplify the analysis in subsequent sections. We approximate (3) by a finite dimensional analogue and obtain a matrix-vector equation that is used in our classification process. Taking the Fourier transform of (3) and assuming a spatially sampled reflectivity function $D(m,n)$ on a grid indexed by m and n we obtain

$$E(\omega) = \sum_n \frac{\Delta(\omega,n)}{n\Delta y} S(\frac{\omega}{n\Delta y}) \quad (4)$$

where

$$\Delta(\omega,n) = \sum_m D(m,n)e^{-j\omega m} \quad (5)$$

ω is the discrete frequency variable and Δy is the sampling interval of y.

The discrete Fourier transform of $E(\omega)$ is obtained by sampling ω at $2\pi/M$ intervals, where M is the number of sampling points. In matrix form (4) then becomes

$$\mathbf{e} = \mathbf{A}\ \mathbf{s} \quad (6)$$

where $\mathbf{A}$ is formed by reordering its elements according to the frequency values at which $S(\frac{2\pi k/M}{n\Delta y})$ is defined $\forall k$ and n and $\mathbf{e}$ is the discrete Fourier transform of the echo signal. The resulting matrix $\mathbf{A}$ has dimension $M \times L$. The vector $\mathbf{s}$ is formed by ordering $S(\frac{2\pi k/M}{n\Delta y})$ with respect to an ascending order of frequency.

3 Target Modeling and Classification

3.1 Target model

We model the reflectivity functions of targets as two-dimensional stochastic processes. There are several reasons for that. The class representative reflectivity function which is either determined by experimental measurements or computation will in general differ from the actual target reflectivity due to variations in relative orientation between the radar and the target. Also, one can view targets as being composed of a small number of strong scatterers and a larger number of weak scatterers whose contribution appears to be random with some correlation structure. A third reason is target fluctuations between looks which introduces randomness in the returns of different transmissions.

Now observe that we use equation (6) to describe target returns. A target is completely specified in our model by a unique matrix $\mathbf{A}$. Since the target reflectivity functions are assumed to be stochastic in nature, the matrices $\mathbf{A}$ in (6) will inherit this property. Target classes are therefore described each by a random matrix with certain statistical properties. The set of waveforms designed to discriminate between classes will naturally be influenced by our choice of statistical model for the target reflectivity function.

3.2 Signal selection strategy

The test that we derive is sequential where the number of transmissions is initially unknown and may vary for each test. The stopping criteria depends on the desired performance. Each target class corresponds to a different statistical hypothesis, i.e.,

$$\begin{aligned} H_\circ : &\quad target\ 0 \\ H_1 : &\quad target\ 1 \end{aligned} \quad (7)$$

We use the *Kullback-Liebler information number* (KLIN) between the probability densities of both hypotheses as a cost function to be optimized to arrive at the discriminating signals. The KLIN when f_1 is the true density is defined by [5]

$$I(f_\circ, f_1) = \int \cdots \int log(\frac{f_\circ(\mathbf{x})}{f_1(\mathbf{x})}) f_\circ(\mathbf{x}) d\mathbf{x} \quad (8)$$

where $f_\circ(\mathbf{x})$ and $f_1(\mathbf{x})$ are the probability densities both hypotheses. The KLIN is the information a random sample brings about hypothesis i if H_i is the true hypothesis.

There are two reasons for using the KLIN to derive a signal selection strategy. For a given probability of false alarm, maximizing the KLIN also asymptotically (i.e., when the number of observations is large) maximizes the probability of detection [5]. Second, when observations are independent, the expected number of observations needed under a specific hypothesis to achieve a desired confidence level is inversely proportional to the KLIN corresponding to that hypothesis as the desired confidence level increases. Therefore, by maximizing the KLIN after each new observation is received we effectively minimize the average number observations needed to reach a decision. Although the observations in our case may not be independent, this result is generally true for correlated Gaussian observations [4]. The KLI numbers assuming $f_\circ$ and f_1 to be the true distributions are in general different.

4 Signal Selection with Gaussian Target Reflectivity Functions of Different Means and Identical Covariances

To illustrate our approach, let us consider in some detail a specific model for the target reflectivity functions under the two hypotheses. In this model, both KLI numbers turn out to be the same. This leads to a fixed waveform selection strategy where the same waveforms are transmitted regardless of which hypotheses is true. Hence, the waveforms can be designed off-line assuming that the intensity of the observation noise is known.

4.1 The Gaussian model

Our model is given by

$$\begin{aligned} H_\circ : &\quad \mathbf{r}_i = (\mathbf{A}_\circ + \mathbf{W})\,\mathbf{s}_i + \mathbf{v}_i \qquad i = 1, 2, \cdots \\ H_1 : &\quad \mathbf{r}_i = (\mathbf{A}_1 + \mathbf{W})\,\mathbf{s}_i + \mathbf{v}_i \qquad i = 1, 2, \cdots . \end{aligned} \quad (9)$$

In (9) each reflectivity matrix is composed of a constant and a random part denoted by $\mathbf{W}$ which we will assume to be Gaussian with uncorrelated zero-mean elements and variance σ_w^2. That is to say that each target is composed of a constant known mean that is either experimentally or theoretically computed and a random part due to aspect uncertainties. The transmitted signals $\mathbf{s}_i$ are assumed to be of unit energy. After each transmission, the vector $\mathbf{r}_i$ is received corrupted by a Gaussian vector $\mathbf{v}$ with covariance matrix $\sum_v$. This vector captures both the receiver/atmospheric noise as well as randomness in target fluctuations. For simplicity we will assume that $\sum_v = \sigma_v^2 \mathbf{I}$ where $\mathbf{I}$ is the identity matrix. Since a hypothesis testing problem is invariant to a bias adding transformation, we can rewrite (9) as

$$\begin{array}{lll} H_\circ: & \mathbf{r}_i = (\mathbf{A}_\circ - \mathbf{A}_1 + \mathbf{W})\,\mathbf{s}_i + \mathbf{v}_i & i = 1, \cdots \\ H_1: & \mathbf{r}_i = \mathbf{W}\,\mathbf{s}_i + \mathbf{v}_i & i = 1, \cdots. \end{array}$$

In a sequential test the number of transmissions needed to reach a decision is initially unknown and is dependent on the amount of information contained in the observations. An attempt to make a decision is made after each return until the terminal condition is satisfied. The waveform transmitted at each stage in the process is determined by the likelihood of each of the targets. At each stage one of two waveforms corresponding to each target is selected depending on which target is more likely. In our present target model, it will be shown that the choice of waveform at each stage is independent of the target type and only depends on the ratio σ_v^2/σ_w^2.

4.2 Sequential test

Let n be the current decision stage and let

$$\mathbf{r} = \left[\begin{array}{cccc} \mathbf{r}_1 & \mathbf{r}_2 & \cdots & \mathbf{r}_n \end{array} \right]^T \tag{10}$$

be the vector composed of all n observations. Under the null hypothesis, $\mathbf{r}$ is a Gaussian random vector of length $(n \times L)$ with mean

$$E[\mathbf{r}|H_\circ] = [\mathbf{I}_n \otimes (\mathbf{A}_\circ - \mathbf{A}_1)] \cdot \mathbf{s} = \mathbf{G} \cdot \mathbf{s} \tag{11}$$

where $\otimes$ denotes the Kronecker product and $\mathbf{s} = [\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_n]^T$. The covariance matrix is given by

$$\mathbf{K}_r = \left[\begin{array}{cccc} \sigma_w^2 + \sigma_v^2 & \sigma_w^2 \cos\theta_{12} & \cdots & \sigma_w^2 \cos\theta_{1n} \\ \sigma_w^2 \cos\theta_{21} & \sigma_w^2 + \sigma_v^2 & \cdots & \sigma_w^2 \cos\theta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_w^2 \cos\theta_{n1} & \sigma_w^2 \cos\theta_{n2} & \cdots & \sigma_w^2 + \sigma_v^2 \end{array} \right] \otimes \mathbf{I}_M$$

where $\cos\theta_{ij} = \mathbf{s}_i^T \mathbf{s}_j$ and M is the row dimension of $\mathbf{A}_\circ$ or $\mathbf{A}_1$. Under the alternative hypotheses, the mean and covariance of $\mathbf{r}$ are $E[\mathbf{r}|H_1] = \mathbf{0}$ and $\mathbf{K}_r$, respectively.

At each decision stage k the observation vector $\mathbf{r}$ is appended with a new observation and a new log-likelihood ratio is computed as follows

$$\begin{aligned} \log \Lambda_k(\mathbf{r}) &= \log\{\frac{f(\mathbf{r}_k|H_1)}{f(\mathbf{r}_k|H_\circ)}\} & (12) \\ &= \mathbf{s}^T \mathbf{G}^T \mathbf{K}_r^{-1} \mathbf{r} - \frac{1}{2}\, \mathbf{s}^T \mathbf{G}^T \mathbf{K}_r^{-1} \mathbf{G}\, \mathbf{s} & (13) \end{aligned}$$

In a sequential test the likelihood ratio at each stage is compared to the thresholds $A > B$ [3]. If it happens that $B < \Lambda_k(\mathbf{r}) < A$ the test is continued. This process terminates at some stage n if either $\Lambda_n(\mathbf{r}) > A$ or $\Lambda_n(\mathbf{r}) < B$. In such a case, we accept $H_\circ$ if $\Lambda_n(\mathbf{r}) > A$ and H_1 if $\Lambda_n(\mathbf{r}) < B$

The motivation for the sequential test described above is to transmit signals only when necessary to make the decision. In some instances the observations point clearly to one particular hypothesis. In these cases, only a few number of transmissions are necessary. This is contrary to the fixed sample size test where the number of observations is always the same. From (13), it is simple to show that the KLIN under both hypotheses is given by

$$KLIN = (1/2)\, \mathbf{s}^T \mathbf{G}^T \mathbf{K}_r^{-1} \mathbf{G}\, \mathbf{s}. \tag{14}$$

Hence, once the signal set is determined, the sequence of their transmission will also be known regardless of which hypothesis is true. This results in a fixed waveform selection strategy.

Note that it may be shown that for the particular model that we have assumed in this section, minimizing the expected number of observations required to reach A or B is equivalent to maximizing (14) for all values of A and B. In other words, we can show that unlike the general case where our approach can be shown to be optimal when the desired confidence levels are high, our approach in this case is optimal regardless of the desired confidence level.

4.3 Solution to the signal selection problem

In order to determine the waveforms we start at the first stage when $n = 1$ and find $\mathbf{s}_1$ that maximizes

$$\mathbf{s}^T \mathbf{G}^T \mathbf{K}_r^{-1} \mathbf{G}\, \mathbf{s} = \mathbf{s}_1 \mathbf{A}^T \mathbf{A} \mathbf{s}_1 / (\sigma_w^2 + \sigma_v^2) \tag{15}$$

where $\mathbf{A} = \mathbf{A}_\circ - \mathbf{A}_1$. The right-hand side followed from the definition of $\mathbf{K}_r$ in (12). The maximizing signal $\mathbf{s}_1^*$ is the eigenvector of $\mathbf{A}^T\mathbf{A}$ corresponding to its maximum eigenvalue λ_1. In general if a decision is not reached after $n-1$ transmissions, the nth waveform $\mathbf{s}_n$ is transmitted. This signal is determined by maximizing $\mathbf{s}^T \mathbf{G}^T \mathbf{K}_r^{-1} \mathbf{G}\, \mathbf{s}$ where $\mathbf{s} = [\mathbf{s}_1^{*T}\ \mathbf{s}_2^{*T}\ \cdots\ \mathbf{s}_{n-1}^{*T}\ \mathbf{s}_n]$ The signals $\mathbf{s}_1^*, \mathbf{s}_2^*, \cdots, \mathbf{s}_{n-1}^*$ are the optimum signals transmitted up to stage $n-1$. Let $\mathbf{s}_1^*, \mathbf{s}_2^*, \cdots, \mathbf{s}_{n-1}^*$ be an orthonormal basis for some subspace V of $\mathcal{R}^L$. One can then write $\mathbf{s}_n$ as

$$\mathbf{s}_n = a_1 \mathbf{s}_1 + a_2 \mathbf{s}_2 + \cdots + a_n \mathbf{s}^\perp \tag{16}$$

where $\mathbf{s}^\perp \in V^\perp$ and $\sum_{i=1}^{n} a_i^2 = 1$. One can show that

$$\begin{aligned} \mathbf{s}^T \mathbf{G}^T \mathbf{K}_r^{-1} \mathbf{G}\, \mathbf{s} &= \frac{\frac{1}{\sigma_w^2}[x^2(\lambda + \sum_{k=1}^{n-1} \lambda_k) + \mathbf{a}^T \mathbf{Q} \mathbf{a}]}{x[x^2 - \mathbf{a}^T \mathbf{a}]} & (17) \\ &= f(\mathbf{a}, \lambda) & (18) \end{aligned}$$

where $\lambda = \mathbf{s}^\perp \mathbf{A}^T \mathbf{A} \mathbf{s}^\perp$, $x = (1 + \sigma_v^2/\sigma_w^2)$, $\mathbf{a} = \left[\begin{array}{cccc} a_1 & a_2 & \cdots & a_{n-1} \end{array} \right]$ and $\mathbf{Q} = [q_{ii}]$ is a diagonal matrix with entries $q_{ii} = x^2(\lambda_i - \lambda) - 2x\lambda_i - \sum_{j \neq i}^{n-1} \lambda_j$.

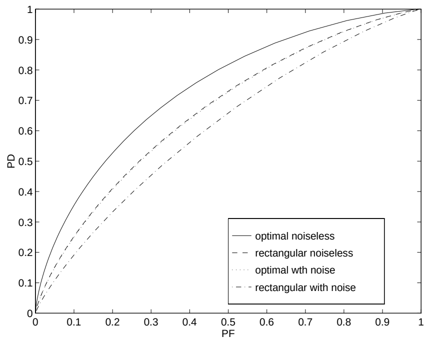

Figure 1: ROC curves for $n = 1$ with optimal and *rect* waveforms.

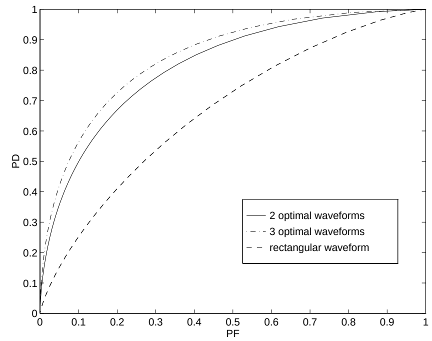

Figure 2: ROC curves for $n = 2$ and $n = 3$ with $\sigma_v^2 = 0$.

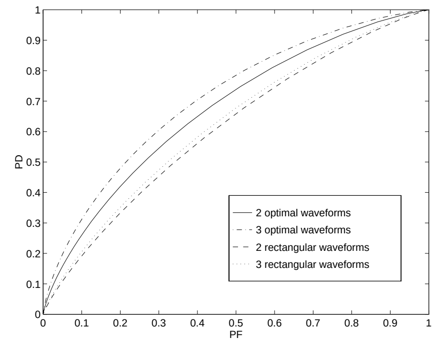

Figure 3: ROC curves for $n = 2$ and $n = 3$ for $\mathbf{s}_2 \perp \mathbf{s}_1$ and $\mathbf{s}_3 = \mathbf{s}_1$

The optimum signal $\mathbf{s}_n$ is found by maximizing $f(\mathbf{a}, \lambda)$ in (18) with respect to $\mathbf{a}$ over the unit sphere $\|\mathbf{a}\| \leq 1$ since $\|\mathbf{a}\|_2^2 = 1 - a_n^2$. The Kuhn-Tucker conditions for this problem result in two possible solutions for $\mathbf{a}$. One solution is at $\mathbf{a} = 0$ while the other is a degenerate case that also includes $\mathbf{a} = 0$ as a solution. The type of the stationary point at $\mathbf{a} = \mathbf{0}$ therefore determines the outgoing signal at the nth stage.

If $\mathbf{a} = \mathbf{0}$ is a maximum, then $a_n = 1$ and $\mathbf{s}_n \in V^\perp$. It can be shown that $\mathbf{a} = \mathbf{0}$ is a maximum if

$$(1 + \sigma_v^2/\sigma_w^2) \leq (1 + \lambda/\lambda_1)/(1 - \lambda/\lambda_1) \qquad (19)$$

and that $f(\mathbf{0}, \lambda)$ is maximum when $\lambda = \lambda_n$, where λ_n is the nth largest eigenvalue of $\mathbf{A}^T\mathbf{A}$. The signal $\mathbf{s}_n$ is therefore the eigenvector of $\mathbf{A}^T\mathbf{A}$ corresponding to λ_n.

If $\mathbf{a} = \mathbf{0}$ is a saddle point or a minimum, the maximum of $f(\mathbf{a}, \lambda)$ occurs on $\|a\| = 1$ and therefore $\mathbf{s}_n \in V$. It is simple to show that

$$arg \max_{\|\mathbf{a}\|=1} f(\mathbf{a}, \lambda) = [1\ 0\ \cdots\ 0]^T \qquad (20)$$

and hence that $\mathbf{s}_n = \mathbf{s}_1$. Therefore, when $\sigma_v^2/\sigma_w^2 > 2(\lambda_n/\lambda_1)/(1 - \lambda_n/\lambda_1)$, more information about the target can be obtained by sending a repetition of the first transmitted signal than any other signal. In this case the following transmissions will have to be $\mathbf{s}_1$ to maximize information about the target.

5 Simulation Results

We simulated a case with 2 targets and obtained the optimal waveforms for different values of the ratio σ_v^2/σ_w^2. The performance of each case is examined by drawing the receiver operating characteristics (ROC) for a given value of n. The probability of false alarm for a given n is defined as $P_F|_n = Prob\{acceptH_\circ|H_1, n\}$ and the probability of detection for a given n is defined as $P_D|_n = Prob\{acceptH_\circ|H_\circ, n\}$.

Figure 1 compares the performance of the optimal waveform to that of an arbitrary waveform, in the presence and absence of an observation noise. The arbitrary waveform was chosen to be a rectangular (*rect*) function. In the noiseless case, the optimal waveform performs much better than the arbitrary signal. When noise is added, the advantage of using the optimal waveform diminishes as σ_v^2 is increased.

Figure 2 shows a noiseless case with $n = 2$ and $n = 3$. The performance of the rectangular waveform does not change in this case as a function of n. This is because when $\sigma_v^2 = 0$, no new information is added by retransmitting the same signal. Improvements are obtained by transmitting the optimal set for this case. The performance is a function of $\sum_{i=1}^n \lambda_i$ as seen from (17) by setting $\mathbf{a} = 0$.

In Fig. 3 noise is added such that the optimal set consists of the first two orthonormal signals for $n = 2$ and then a repetition of the first signal when $n = 3$. Obviously, the performance of the optimal waveforms deteriorates with the addition of noise as compared to Fig. 2.

References

[1] E. J. Kelly and R.P. Wishner, "Matched-Filter Theory for High-Velocity Accelerating Targets," *IEEE Trans. Military Electronics*, vol. MIL-9, Jan. 1965, pp. 56-69.

[2] N.F. Ezquerra, "Target Recognition Considerations," in *Principles of Modern Radar*, J. L. Eaves and E.K. Reedy, eds., New York: Van Nostrand Reinhold, 1987.

[3] A. Wald, *Sequential Analysis*, John Wiley & Sons, New York, 1947.

[4] J. J. Bussgang and D. Middleton, "Optimum Sequential Detection of Signals in Noise," *IRE Trans. on Inform. Theory*, vol. IT-1, pp. 5-18, 1955.

[5] H. Chernoff, *Sequential Analysis and Optimal Design*, Society for Industrial and Applied Mathematics, Philadelphia, 1972.

[6] C.W. Therrien, "A Sequential Approach to Target Discrimination," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-14, no.3, pp. 433-440, May 1978.