Optimizing decision threshold and weighting parameter for UXO discrimination

Haoping Huang¹, Bill SanFilipo², and I. J. Won²

ABSTRACT
Establishing a decision threshold to separate unexploded ordnance (UXO) from clutter is a critical task in UXO discrimination procedures. It is, however, a difficult task because the decision threshold depends on many parameters such as local geology, distribution of UXO and clutter, and cultural and environmental noise at a given UXO cleanup site. Some discrimination algorithms also use site-dependent weighting factors. In practice, one determines these parameters empirically from calibration experiments. To facilitate this task, we introduce a simple function called the degree of discrimination, constructed from calibration plots at a UXO cleanup site. First, we compute the degree of discrimination for a representative set of thresholds and weighting factors using ground truth data from the calibration site. Then, we find the optimum decision threshold and weighting factors associated with the best degree of discrimination. Finally, we use these optimal parameters for the UXO site cleanup. We have conducted two experiments using the broadband electromagnetic (EM) sensor with a concentric coil configuration. Our experiments show that 92% of UXO is declared and 74% of clutter is rejected using our optimized threshold and weights. A blind test shows that optimizing the threshold and weights increases the degree of discrimination from 0 to 0.15. Our tests also show that the detection scheme with a collective threshold for all UXO types is easier to use and yields better discrimination than the scheme based on individual thresholds for each UXO type.

INTRODUCTION
Most metal detectors can detect metal pieces but cannot effectively discriminate unexploded ordnance (UXO) in a cluttered environment. False alarms are anomalies caused by other metallic objects, soil heterogeneities, and other natural and cultural features. False alarms result in unnecessary excavations and therefore contribute significantly to the cost of UXO clearance. A major research initiative is to develop discrimination (target identification) capabilities. Perhaps the most promising subsurface sensing technologies for UXO discrimination are electromagnetic (EM) methods (e.g., Won et al., 1998; Paison et al., 2001; Morrison et al., 2004; Tarokh et al., 2004; O’Neill et al., 2006). Each UXO type generates a characteristic EM response that is a function of its size, shape, depth, orientation, conductivity, and magnetic permeability. Even though combinations of these factors may yield similar EM responses for different objects, these approaches can reduce the number of false alarms.

A general approach for identifying a UXO item, as well as for discriminating between UXO and clutter, is to match the EM response at multiple frequencies or time channels measured above a target against signatures from a set of known UXO types that are stored in a library. If the misfit is below a given threshold, the object may be accepted as a corresponding UXO; otherwise, it is declared as clutter. Therefore, establishing a decision threshold is critical but difficult in target discrimination. Many algorithms have been developed to discriminate UXO from clutter. Some require accurate knowledge of the relative sensor/target positions from spatial data (Bell et al., 1998; Ozdemir et al., 1998; Bell et al., 2001) and some do not (Norton et al., 2001). However, all algorithms need a decision threshold of some kind to separate UXO from clutter. Unfortunately, such a threshold can only be established from experiments because it depends on the distribution of UXO and clutter, local geology, and cultural and environmental noise levels of each cleanup site. Under most circumstances, a simple threshold is selected subjectively, based on experience. By comparing the performance of several discriminating algorithms, Collins et al. (2001) have found that a simple threshold on EM data is not an effective discriminator of UXO from clutter items.

The EM data from UXO and clutter items have overlapping distributions. As a result, there is always a trade-off between the UXO
declaration and the false alarm rate. Usually, the false alarm rate increases with an increase in detection rate. We seek an approach that can be used to select a cut-off point to yield the maximum UXO detection with minimum false alarms. This is not a trivial problem. One way to do this is to use the features of the receiver operating characteristic (ROC) curves. These curves are used to evaluate the results of a classification process and were first employed in the study of discriminator systems for detecting radio signals in the presence of noise in the 1940s. In the 1960s, they began to be used in psychophysics to assess human (and occasionally animal) detection of weak signals. They are also used extensively in making medical decisions (e.g., Swets, 1995) as well as in evaluating machine-learning results.

The ROC curves are an appropriate tool to provide a clear measure of the overall performance of a sensor and therefore are widely used to evaluate the sensors and algorithms applying to detection and discrimination of UXO and landmines (e.g., Collins et al., 2001; Asch et al., 2002; Zoubir et al., 2002). However, they have rarely been used in optimizing the decision threshold to distinguish UXO/landmines from clutter or to optimize parameters in discriminating algorithms. Bennett et al. (1999) and Haskett and Rupp (2002) study airborne laser imagery data used in UXO/landmine detections. The former presents the enhanced performance of prior classification work of airborne imagery data by introducing a rejection class, which accurately describes a particular source of false alarm. This enhancement is measured by observing changes in the ROC curves corresponding to two different fixed strengths of the rejection class. The latter indicates that the ROC curves provide a good basis for system trade-off studies. Müller et al. (2002) discuss the reliability of mine detection systems and strongly recommend applying the ROC curves to the assessment and optimization of mine detecting systems. Saito et al. (2002a, b) generate a hypothetical UXO site using stochastic simulation and then examine, by mimicking the use of the ROC curve, the impact of the quality of prior information available and the choice of an acceptable false alarm rate on the statistical characterization of UXO.

In this paper, we discuss how to optimize the decision threshold and weighting factor using ROC curves obtained from the calibration or sensor training plot at a cleanup site. As an example, we use a handheld broadband EM sensor and a simple discrimination approach based on a linear combination of the target’s longitudinal and transverse responses (Won et al., 1997; Norton et al., 2001).

### THE EM SENSORS

The GEM-3 broadband electromagnetic sensor has been used at many environmental sites, including those containing landmines and UXO (Gao et al., 2000; Won et al., 2001; Collins et al., 2002; Nelson et al., 2002; Huang and Won, 2003a; Nelson et al., 2005). The sensor (Figure 1), designed particularly for detecting small metallic targets, can be handheld or cart mounted (Keiswetter et al., 1997; Won et al., 1997).

The sensor operates in a bandwidth from 30 Hz to 96 kHz. The sensing head consists of a pair of concentric, circular coils (Tx1 and Tx2) that transmit a continuous, broadband, digitally controlled EM waveform. The two transmitter coils with precise dimensions and placement are connected in an opposing polarity and create a magnetic cavity at the center of the two coils. A magnetic cavity is defined as a region where a directional sensor, placed in a specified orientation, produces zero signal induced from the magnetic field. The receiving coil (Rx) is placed within this magnetic cavity so that it senses only the weak secondary field returned from the earth and buried targets.

### LINEAR COMBINATION EM SPECTRA

We use a simple identification approach based on a linear combination of the target’s longitudinal and transverse responses — a procedure that does not require accurate knowledge of the relative sensor/target position, provided only that the sensor/target separation is small enough to produce an adequate target signal. We shall briefly discuss the technique below; the details may be found in Norton et al. (2001). The targets of interest are assumed to be axially symmetric, which, to a reasonable approximation (e.g., ignoring fins), is true of almost all UXO. In the far field, the EM response of an axially symmetric target is entirely characterized by its longitudinal and transverse spectral responses (Das et al., 1990). A dipole approximation of secondary fields induced by the target is quite good when the target/sensor separation is at least three or four times the target size. The longitudinal and transverse responses correspond, respectively, to sensor transmitter primary fields parallel and perpendicular to the target axis.

For UXO at an arbitrary orientation, the recorded spectrum can be expressed as a linear combination of its longitudinal and transverse responses. Our procedure is to determine the linear combination that provides the best fit to the measured response. Once a library of a set of UXO signatures for the longitudinal and transverse responses, \( \beta_l(f_k) \) and \( \beta_t(f_k) \), has been collected and stored, we wish to minimize the misfit,

\[
\chi = \sum_{k=1}^{N} \frac{|V(f_k) - s_L \beta_L(f_k) - s_T \beta_T(f_k)|^2}{w + \sum_{k=1}^{N} |V(f_k)|},
\]

where \( N \) is the number of frequencies, \( f_k \) is the \( k \)th frequency, \( V(f_k) \) is a measurement at the \( k \)th frequency over an unknown target, and \( w \) is a positive weighting factor that may account for differences in the signal-to-noise ratio of the data (SanFilipo et al., 2005). The coefficients \( s_L \) and \( s_T \) depend upon the primary fields and their angles of incidence but are assumed independent of frequency. They are computed for every library spectrum by selecting the coefficients that

---

**Figure 1.** Coil configuration of the broadband EM sensor.
minimize $\chi$ for a given observed spectrum. The library spectrum that produces the smallest mean-square error $\chi$ is then used to identify the most likely target.

## OPTIMIZATION OF DECISION THRESHOLD AND WEIGHT

In general, the threshold $t$ varies depending on UXO type, local geology, clutter density, and cultural and environmental noise levels of each site. It should be established on a site-by-site basis and sometimes on a UXO-type basis as well. Thus, the threshold can be either collective or individual. A collective threshold is that established at a site for all types of UXO stored in the library, and an individual threshold is for one type of UXO. In most circumstances, the signal levels for different targets vary over several orders of magnitude. Thus, the weight $w$ also needs to be optimized to accommodate targets having small signals if a collective threshold is used. Otherwise, individual thresholds must be used for each type of UXO, based on local conditions. Because the thresholds and weighting factor must be obtained from experiments, the sensor training or calibration plot at a UXO cleanup site may be used to establish them.

To establish the library, we measure the longitudinal and transverse responses of a set of known targets. This is done by placing the sensor over each target in the air and recording its response while the target is in the vertical and horizontal orientations relative to the sensor axis. Then, a survey over the calibration plot is performed as practically as possible, and the discrimination process is carried out using various $t$ and $w$. Finally, analysis of the results of discrimination against the ground truth will yield the optimum threshold and weighting factor.

To analyze the discrimination results, we define the following parameters:

- $O_o$: the true positive fraction, i.e., the number of UXO items declared as UXO divided by the number of emplaced UXO items
- $C_a$: the false positive fraction, i.e., the number of clutter items declared as UXO divided by the number of emplaced clutter items
- $O = 1 - O_o$: the false negative fraction, i.e., the fraction of UXO classified as clutter
- $C_c = 1 - C_a$: the true negative fraction.

Each parameter varies between 0 and 1 and depends on $t$ and $w$ for a given data set obtained at a specific site. Since we wish to have $O_o$ and $C_c$ be as large as possible and $C_a$ and $O_c$ as small as possible, maximizing the function $g_{opt}(t,w) = (O_o + C_c)/(C_a + O_c)$ would yield the best discrimination. Considering $O_c$ and $C_a$ are not independent variables, the degree of discrimination can be defined as

$$
\varphi(t,w) = O_o - C_a.
$$

The optimum threshold $t$ and weight $w$ would yield a maximum of $\varphi(t,w)$, i.e., a maximum number of UXO classified with minimum false alarms. The ROC curve is a plot of $O_o$ versus $C_a$, varying the thresholds for a given $w$ in this particular case. Figure 2 illustrates two examples of the ROC curve and $\varphi(t,w)$. It is well known that the closer to the upper-left corner the position of a ROC point is, the better the sensor or algorithm. Figure 2b indicates the best discrimination, $O_o = 0.92$, $C_a = 0.4$, and so $\varphi(t,w) = 0.52$, obtained at $t = 20\%$ for the solid curve; while $O_o = 0.65$, $C_a = 0.5$, and so $\varphi(t,w) = 0.15$ at $t = 25\%$ for the dashed curve. These points are shown by solid circles in Figure 2.

At a threshold and weight of interest, $\varphi(t,w)$ measures the degree to which a sensor or algorithm discrimination function performs. Therefore, we refer to it as the degree of discrimination, which ranges between $-1$ and 1. The best possible discrimination method would yield $\varphi(t,w)$ of one, i.e., all UXO items are found at zero false alarm rate. This is a point in the upper-left corner of the ROC space. A completely random discriminator would give $\varphi(t,w) = 0$, a straight line at an angle of $45^\circ$ from the $O_o$ axis, from bottom left to top right; this is because as the threshold is raised, equal fractions of true and false positives would be let in. Results below this line would indicate a detector giving wrong results consistently; inverting the results would identify detectors that gave useful results.

There are several ways to summarize the ROC curve into a single number. The most common one used in engineering is the area $A$ between the ROC curve and the no discrimination line because of its useful mathematical properties as a nonparametric statistic. This area is often known as the discrimination (strictly, total degree of discrimination). For example, for two ROC curves in Figure 2a, we have $A = 0.04$ and 0.08, respectively. However, it should be kept in mind that any attempt to summarize the ROC curve into a single number loses information about the pattern of trade-offs of the particular discriminator algorithm.

## EXPERIMENT AT U. S. ARMY TEST SITE

We performed an experiment at the Standard UXO Technology Demonstration Site at Aberdeen Proving Grounds (APG), Maryland. This seeded site for controlled testing included (1) calibration lanes for system training and target characterization; (2) blind grid — a 1600 m$^2$ rectangular grid including access lanes separating 400 discrete 1 × 1 m square interrogation points; (3) open area — a large area including some moderately rough terrain; (4) moguls — an area with moguls and craters of about ±1 m vertical relief, requiring manual data acquisition, likely in a hand-held sensor configuration; and (5) wooded area with dense vegetation and wetlands.

The calibration lanes have 195 squares of 1 × 1 m, with 128 seeded squares, including 40 clutter items and 88 UXO objects from 14 types (see Table 1 for the list). The calibration lanes contained 8 mines that were not included in our library and thus will be regarded as UXO.
as clutter in the discrimination process. The library data were collected from the UXO samples in the air at 10 frequencies ranging from 90 Hz to 41 kHz. Figure 3 shows two example spectra obtained for the library.

The EM data were collected over the 195 squares using the same frequencies as those used for the library data. In total, 111 targets were detected, including 75 UXO and 34 clutter objects, with 13 UXO and six clutter items missing. Most missed items were buried deeper than 1 m. In the following discussions, we consider only the targets that are detected, because the topic of this paper is the discrimination problem. Thus, the total numbers are 75 for UXO and 34 for clutter items.

First, we attempt to determine the optimum collective threshold $t$ by running the discrimination procedure using various values of $t$ and $w$. Figure 4 shows the resulting $\varphi(t,w)$ as a 2D image. For this specific distribution of UXO and clutter, the maximum of $\varphi(t,w)$ appears in the range 1%–2% for $t$ and 150–160 for $w$. We choose $t = 1.6\%$ and $w = 155$ based on the image, which should yield the best discrimination results. Figure 5 illustrates the $O_o$, $C_c$, $O_c$, and $C_o$, $\varphi(t,w = 155)$, the percentage of UXO correctly identified by name as a function of threshold for $w = 155$, and the ROC curve. As can be seen in Figure 5a, all objects, either UXO or clutter, are declared as clutter when the threshold is very low (0.1%), while all are declared as UXO when the threshold is very high (30%). The degree of discrimination $\varphi(t,w)$ reaches its maximum at $t = 1.6\%$, where 92% of the UXO are classified as such (8% of UXO declared as clutter) and 74% of the clutter items are rejected, leaving 26% of clutter declared as UXO. If we prefer not to miss any UXO, the threshold can be slightly increased, for example, to 2.5%. As shown in Figure 5a, all UXO items (75) are correctly declared. This results in the false alarm rate being increased from 26% to 44%. The ROC curve and the degree of discrimination $\varphi(t,w = 155)$ are denoted in Figure 5b and c. As can be seen, the peak of $\varphi(t,w)$ is very distinct at $t = 1.6\%$, indicating that the optimized threshold is much better than the others. Figure 5d indicates that 47% of the UXO items were correctly identified by name at $t = 1.6\%$. Table 2 shows the details on the optimum collective threshold and weight.

Next, we examine the individual thresholds. The weight $w$ in equation 1 is ignored, and the discrimination procedure is run for various values of $t$ to determine the optimum threshold for each type

Table 1. UXO types buried in the calibration lanes.

<table>
<thead>
<tr>
<th>Type</th>
<th>Number</th>
<th>Description</th>
<th>Depth (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 mmP</td>
<td>6</td>
<td>20-mm projectile</td>
<td>0.1–0.2</td>
</tr>
<tr>
<td>BLU26</td>
<td>6</td>
<td>BLU-26 submunition</td>
<td>0.1–0.2</td>
</tr>
<tr>
<td>M42</td>
<td>6</td>
<td>M42 submunition</td>
<td>0.15–0.35</td>
</tr>
<tr>
<td>40 mmG</td>
<td>6</td>
<td>40-mm grenades</td>
<td>0.1–0.2</td>
</tr>
<tr>
<td>40 mmP</td>
<td>6</td>
<td>40-mm projectile</td>
<td>0.3–0.6</td>
</tr>
<tr>
<td>BDU28</td>
<td>6</td>
<td>BDU-28 submunition</td>
<td>0.1–0.2</td>
</tr>
<tr>
<td>57 mmP</td>
<td>6</td>
<td>57-mm projectile APC M86</td>
<td>0.4–0.91</td>
</tr>
<tr>
<td>60 mmM</td>
<td>10</td>
<td>60-mm mortar</td>
<td>0.25–1.0</td>
</tr>
<tr>
<td>MK118</td>
<td>6</td>
<td>MK118 Rockeye</td>
<td>0.3–0.6</td>
</tr>
<tr>
<td>81 mmM</td>
<td>6</td>
<td>81-mm mortar</td>
<td>0.5–1.5</td>
</tr>
<tr>
<td>2.75 in</td>
<td>6</td>
<td>2.75-inch rocket M230</td>
<td>0.5–1.2</td>
</tr>
<tr>
<td>105H</td>
<td>6</td>
<td>105-mm heat rounds M456</td>
<td>0.4–0.8</td>
</tr>
<tr>
<td>105 mmP</td>
<td>6</td>
<td>105-mm projectile M60</td>
<td>0.9–1.8</td>
</tr>
<tr>
<td>155 mmP</td>
<td>6</td>
<td>155-mm projectile M483A1</td>
<td>0.9–2.0</td>
</tr>
</tbody>
</table>

Figure 3. The longitudinal and transverse spectral responses of two types of UXO: (a, b) 20 mmP, and (c, d) 155 mmP. Solid circles stand for the in-phase component $I$, and open circles for the quadrature component $Q$.

Figure 4. Image of $\varphi(t,w)$ as functions of threshold $t$ and weight $w$.

Figure 5. (a) The $O_o$, $C_c$, $O_c$, and $C_o$, (b) the ROC curve, (c) $\varphi(t,w)$, and (d) the number of UXO correctly identified by name, $N_n$, for given weight $w = 155$. We choose $t = 1.6\%$ and $w = 155$ based on the image, which should yield the best discrimination results.
of UXO. For example, Figure 6 shows the results for two UXO items, 40 mmP and 57 mmP. As can be seen, the optimum thresholds are 20% for 40 mmP and 22% for 57 mmP. Table 3 lists the optimized threshold for each target and compares them with the collective threshold. Both UXO classification and clutter rejection are better for the collective than for the individuals, especially the clutter rejection, which is 33% better.

It would be simpler, in practice, to use the collective threshold rather than the individual threshold. The former can be easily adjusted to satisfy the user’s requirements. For example, the optimized parameters based on the calibration plot may not be perfect for the blind ground, even though the calibration plot is designed to simulate reality as much as possible. Thus, prior biased information may be used, i.e., the threshold is slightly adjusted to yield a reasonable ratio of UXO to clutter. It is obvious that this adjustment is much easier for the collective threshold than for the individual.

The degree of discrimination \( \phi(t, w) \) depends upon the site and distribution of UXO and clutter, so the optimal \( t \) and \( w \) established from one site may not be optimal for other sites.

### RESULTS FROM BLIND TESTS

#### The blind grid

We collected the blind grid data in 2003 at 10 frequencies from 90 Hz to 40 kHz, using a cart-mounted sensor. The testing agency evaluated the blind test results. Based on the threshold we used, 70% of UXO was declared as UXO at a false alarm rate of 70%, yielding a degree of discrimination of 0. The ROC curve constructed by the agency is close to the no-discrimination line.

The ground truth of the blind grid has been released recently. Because we did not survey the calibration lanes in 2003, we used the data from the blind grid and the ground truth to determine optimum threshold and weighting factor. The resulting \( \phi(t, w) \) is illustrated in Figure 7, in which we see the maximum of \( \phi(t, w) \) located at \( t = 7.3\% \) and \( w = 35 \). Figure 8 shows the \( O_o, C_c, O_c, \) and \( C_o, \phi(t, w = 35) \), and the percentage of UXO correctly identified by name as a function of threshold at a specified weight \( w = 35 \). The curves of \( O_o \) and \( C_o \) (and so \( C_c \) and \( O_c \)) are very close to each other, indicating poor discrimination (Figure 8a). We reprocessed the data using the optimum collective threshold \( t = 7.3\% \) and weight \( w = 35 \). As shown by the ROC curve in Figure 8b, we have 79% of the UXO correctly classified at a false alarm rate of 64%, i.e., the UXO detection rate is increased by 9% and false alarm rate is decreased by 6%. The degree of discrimination increases from 0 to 0.15. However, only 18% of UXO are identified correctly by name, as shown in Figure 8d.

#### Geophex test site

Our second example is from Geophex’s site in Raleigh, North Carolina. The \( 10 \times 10 \text{ m} \) site contains a total of 21 metal pipes of various diameters and lengths (Huang and Won, 2003a, b). A cart-mounted sensor with differential GPS was used to collect the EM

<table>
<thead>
<tr>
<th>True</th>
<th>20 mmP</th>
<th>BLU26</th>
<th>M42</th>
<th>40 mmG</th>
<th>40 mmP</th>
<th>BDU28</th>
<th>57 mmP</th>
<th>60 mmM</th>
<th>MK118</th>
<th>81 mmM</th>
<th>2.75 in</th>
<th>105H</th>
<th>105 mmP</th>
<th>155 mmP</th>
<th>Clutter</th>
<th>Missed</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 mmP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLU26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M42</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40 mmG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40 mmP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BDU28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>57 mmP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60 mmM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MK118</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>81 mmM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.75 in</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>105H</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>105 mmP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>155 mmP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clutter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td>10</td>
<td>10</td>
<td>8</td>
<td>6</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>31</td>
<td>19</td>
<td>128</td>
<td></td>
</tr>
</tbody>
</table>

### Table 3. Comparison of the discrimination results from the collective and individual thresholds. The numbers are in percent.

<table>
<thead>
<tr>
<th>Threshold</th>
<th>( O_o )</th>
<th>( C_c )</th>
<th>( O_c )</th>
<th>( C_o )</th>
<th>( N_n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collective</td>
<td>0.92</td>
<td>0.74</td>
<td>0.08</td>
<td>0.26</td>
<td>0.47</td>
</tr>
<tr>
<td>Individual</td>
<td>0.89</td>
<td>0.41</td>
<td>0.11</td>
<td>0.59</td>
<td>0.45</td>
</tr>
</tbody>
</table>
data at 10 frequencies (90 Hz–48 kHz). The survey includes the seeded test site and the surrounding area where many known and unknown clutter items such as a manhole cover, sprinklers, and metal debris exist.

We picked 88 anomalies based on the method described in Huang and Won (2003b). The southern part of the site was used as a calibration ground, which contains all seeded pipes and 30 clutter items (Figure 11). The degree of discrimination \( \phi(t, w) \) is illustrated in Figure 9. The high value feature is very sharp, indicating that degree of discrimination depends strongly upon the threshold and weight. The maximum of \( \phi(t, w) \) occurs at \( t = 4.7\% \) and \( w = 27 \). Figure 10a and b shows the ROC curve and \( \phi(t, w = 27) \) for the calibration ground; 95% of the pipes were correctly declared, with a false alarm rate of 43%. We applied the optimum collective threshold \( t = 4.7\% \) and weight \( w = 27 \) to the whole area. As shown by the ROC curve in Figure 10c, the discrimination rate remains 95%, with a false alarm rate of 37%, i.e., 63% of clutter rejected. The degree of discrimination is 0.58, as shown in Figure 10d. There are 12 (57%) pipes identified correctly by name. Figure 11 illustrates a map of the total apparent conductivity on which the results of target classification are posted (Huang and Won, 2004). As can be seen, only one of the pipes is declared as clutter. Actually, it best matches the correct item, but the misfit is higher than the optimum threshold \( t = 4.7\% \).

---

Figure 6. The \( \phi_o, \phi_c, \phi_e \) and \( \phi_o \) as a function of threshold, (a) and (b) for 40 mmP and (c) and (d) for 57 mmP.

Figure 7. Image of \( \phi(t, w) \) as functions of threshold \( t \) and weight \( w \) for the blind grid.

Figure 8. (a) The \( \phi_o, \phi_c, \phi_e \) and \( \phi_o \), (b) the ROC curve, (c) \( \phi(t, w) \), and (d) the number of UXO correctly identified by name for \( w = 35 \).

Figure 9. Image of \( \phi(t, w) \) as functions of threshold \( t \) and weight \( w \) for the assumed calibration ground.

Figure 10. The ROC curves and \( \phi(t, w) \) (a, b) for the assumed calibration ground and (c, d) for the whole area.
CONCLUSIONS

A decision threshold used to separate UXO from clutter depends upon local geology, distribution of UXO and clutter, and cultural and environmental noise at a given UXO cleanup site. Establishing such a threshold is a critical and difficult part in a UXO discrimination procedure. We define a simple function, the degree of discrimination, based on the ROC constructed from the data obtained over a calibration plot at a UXO cleanup site. This function is used to optimize the decision thresholds and weighting parameter, which are then used in the discrimination procedure for data obtained in the blind area. Our experiments show that the optimum threshold and weighting factor yield better discrimination results and that the collective threshold, i.e., a single one for all UXO types, is, in practice, easier to use than the individual thresholds, i.e., one for each UXO type.

Because the threshold and weight must be developed for each new UXO clearance operation, this approach is highly contingent upon a priori knowledge of the distribution of UXO and clutter in the calibration site. Also, success depends upon the similarity between the calibration site and the real sites. If a calibration plot is not available, we should conservatively excavate as many targets as practical in the calibration site and the real sites. If a calibration plot is not available, our experiments show that the optimum threshold and weighting factor can be obtained using the method described above.

ACKNOWLEDGMENTS

This study has been partially funded by the Department of Defense Strategic Environmental Research and Development Program (SERDP) and Environmental Security Technology Certification Program (ESTCP) in Arlington, Virginia.

REFERENCES


——–, 2002b, Accounting for uncertainty attached to ROC curves in geostatistical characterization of UXO sites, The 2002 SERDP and ESTCP.
Partners in Environmental Technology Technical Symposium & Workshop.