Landmine Detection and Discrimination Using High-Pressure Waterjets

Daryl G. Beetner

Electrical and Computer Engineering, University of Missouri-Rolla, Rolla, MO 65409, USA
Email: daryl@umr.edu

R. Joe Stanley

Electrical and Computer Engineering, University of Missouri-Rolla, Rolla, MO 65409, USA
Email: stanleyr@umr.edu

Sanjeev Agarwal

Electrical and Computer Engineering, University of Missouri-Rolla, Rolla, MO 65409, USA
Email: sanjeev@umr.edu

Deepak R. Somasundaram

Electrical and Computer Engineering, University of Missouri-Rolla, Rolla, MO 65409, USA
Email: drsz9f@umr.edu

Kopal Nema

Electrical and Computer Engineering, University of Missouri-Rolla, Rolla, MO 65409, USA
Email: ksnty5@umr.edu

Bhargav Mantha

Electrical and Computer Engineering, University of Missouri-Rolla, Rolla, MO 65409, USA
Email: bsmzpd@umr.edu

Methods of locating and identifying buried landmines using high-pressure waterjets were investigated. Methods were based on the sound produced when the waterjet strikes a buried object. Three classification techniques were studied, based on temporal, spectral, and a combination of temporal and spectral approaches using weighted density distribution functions, a maximum likelihood approach, and hidden Markov models, respectively. Methods were tested with laboratory data from low-metal content simulants and with field data from inert real landmines. Results show that the sound made when the waterjet hit a buried object could be classified with a 90% detection rate and an 18% false alarm rate. In a blind field test using 3 types of harmless objects and 7 types of landmines, buried objects could be accurately classified as harmful or harmless 60%–90% of the time. High-pressure waterjets may serve as a useful companion to conventional detection and classification methods.

Keywords and phrases: signal processing, classification, pattern recognition, high-pressure waterjet, object detection, unexploded ordnance.

1. INTRODUCTION

The United Nations estimates that millions of mines lie buried around the world. Improving landmine detection capability is paramount to saving lives of innocent victims. There are numerous landmine detection systems under investigation, including thermal, chemical, acoustic, hyperspectral imagery, ground penetrating radar (GPR), and metal detectors (MD) [1, 2, 3, 4, 5]. Only a few are actively used in the field. Hand-held units utilizing MDs are commonly used. Landmine metal content, soil conditions, and depth are particularly relevant for the MD. Size and shape of the buried object, soil conditions, mine burial depth, and object similarity to landmines provide constraints for MD- and GPR-based landmine detection capability [6, 7, 8]. MDs have proven successful with metallic-based landmines. However, there are many landmines that are plastic-cased and contain minute amounts of metal. The MD responses for these landmine
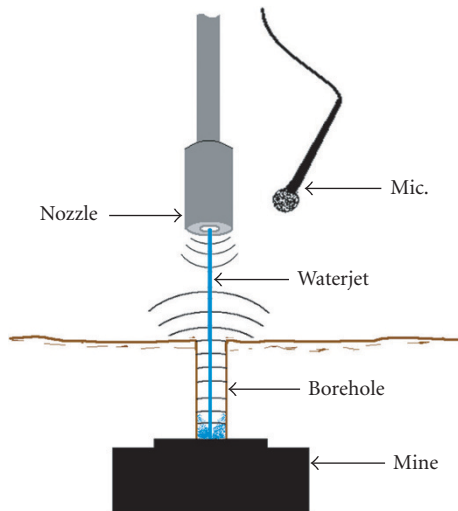
FIGURE 1: A high-pressure waterjet rapidly bores a hole through the soil to strike a buried object. The impact of the waterjet with the buried object creates sounds which are indicative of that object. A typical antipersonnel mine may be 3″ in diameter and buried 2″ deep. The microphone and nozzle are typically located 1″–4″ above the soil surface.

types are often weak, making it difficult to differentiate the plastic landmines from the mineral content of the surrounding soil. Due to high sensitivity, an MD very often provides a false positive signal for small metal debris. GPR sensors have proven more successful in detecting plastic-cased mines. However, GPR sensor systems often suffer from high false-alarm rates since they respond to dielectric discontinuities in metallic and nonmetallic objects. As a result, there is a need for confirmation sensors to help resolve false alarms. Furthermore, the MD- and GPR-based systems provide only an approximate location for the potential landmines. A confirmation sensor such as a metal rod is currently used to precisely locate the mine. In this paper, waterjet technology is investigated as a confirmation sensor for landmine location and discrimination.

A high-pressure waterjet, fired at soil, will quickly create a borehole in the soil (Figure 1). If the waterjet hits an object, the object vibrates, producing a sound that may be used to detect and even identify that object [9, 10]. This sound is a function of the waterjet, its angle with respect to the object, the position at which the object is struck, the characteristics of the surrounding environment (soil cover), and the physical characteristics of the object like its shape, elasticity, and mass. The majority of energy in the sound is typically in the range of 2–10 kHz. The total force applied to the object is small, less than 5 pounds for a waterjet fired at 2500 psi through a 0.05″ nozzle. This force is typically much less than what is required to set off a landmine. If needed, even less force can be used by decreasing the pressure or nozzle size. Depending on pressure, nozzle diameter and firing time, the waterjet can penetrate up to 12″ deep [11]. This research in waterjet-based landmine detection is based on the premise that the acoustic signal produced by the impingent waterjet

is characteristically different for different types or classes of objects [9, 10]. Our objective is to show the potential of using the sound produced by a high-pressure waterjet impact to detect and classify buried landmines.

Three methods of detecting and classifying a buried object using the sound of a waterjet impact were investigated. The methods were based on (a) using unique features computed from the correlation of the recorded sound over time with weighted density distribution (WDD) functions, (b) using a maximum likelihood (ML) estimator applied to the power spectral density of the recorded signal, and (c) using a hidden Markov model (HMM) and cepstral coefficients to model the system as a time-dependent random process whose spectral characteristics are governed by a first-order Markov process. A variety of methods to improve the accuracy of these techniques were explored. The theory and rationale behind each of these three methods and their ability to classify objects are summarized in the following sections.

2. THEORY

2.1. Basis functions applied to temporal acoustic data

The first approach investigated computed temporal features of the acoustic signal. To quantify the change in acoustic signal magnitude over time, correlation of the acoustic signal magnitude with a set of basis functions was examined. WDD functions have been applied for computing spatially and temporally distributed features in hand-held units for landmine/no-landmine discrimination from MD signals [12, 13, 14]. Here, we extend this research to the application of the WDD functions for determining temporal features from the magnitude response of an acquired acoustic signal. The application of the WDD functions to waterjet data is intended to quantify two components of the temporal acoustic signal: (1) low frequency content of the acoustic signal and (2) consistency of the acoustic signal magnitude variation for different object types over the duration of the acoustic response. The temporal features are point-to-point correlations of the WDD functions with the sample-by-sample magnitude of the acoustic signal.

Figure 2 shows the WDD functions, W_k (for $k = 1, \ldots, 6$), that were correlated with measured and windowed sound signals. From Figure 2, the WDD function number is given in parentheses. Let $r[n]$ represent the windowed sound signal with N total samples ($n = 1, \ldots, N$). The WDD functions are piecewise linear, where the WDD function values for each piecewise linear segment are adjusted based on the number of samples (N) to facilitate point-to-point correlation. Let $W_k[n]$ denote the value of the WDD function at sample position n. Six WDD features, ($f_1, \ldots, f_6$), are computed as

$$f_k = \sum_{i=1}^{N} r[i] W_k[i] \tag{1}$$

for $k = 1, 2, \ldots, 6$. Six additional features, ($f_7, \ldots, f_{12}$), are computed from the absolute difference between consecutive
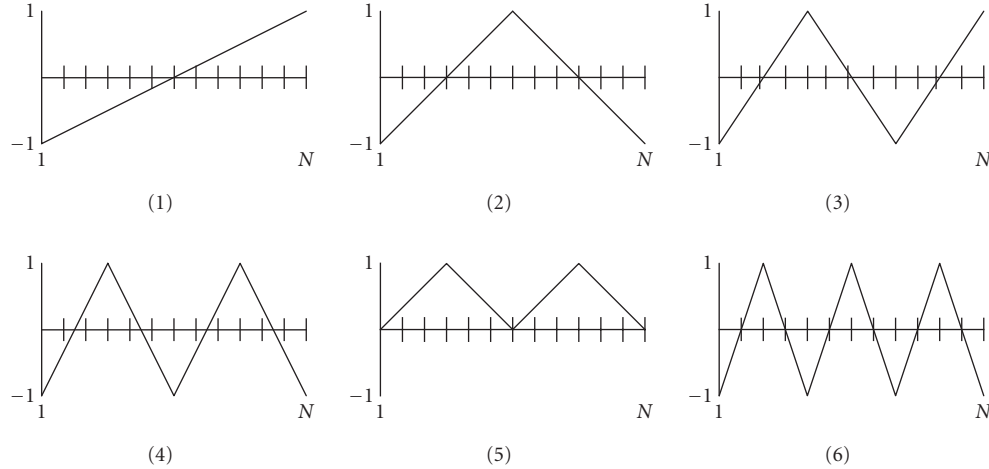
FIGURE 2: WDD functions were correlated with acoustic data produced by the waterjet-mine interaction to calculate temporal features of the acoustic data.

sound values as

$$f_k = \sum_{i=1}^{N} |r[i] - r[i-1]| W_k[i] \qquad (2)$$

for $k = 7, 8, \ldots, 12$, where $r[0] = 0$.

A clustering-based approach was used to discriminate landmines from soil or harmless objects using the twelve WDD features. To compute clusters, the sound data collected at each test site was divided into 10 randomly chosen training and test sets, using 80% of the data for training and the remaining 20% for test (see following sections). K-means clustering [15] of the landmine encounters from the training data was performed to generate a model representation of landmines. The number of clusters, m, was determined empirically.

The nearest neighbor approach was used for landmine discrimination [15]. Let D_i denote the Euclidean distance from cluster i $(1 \le i \le m)$, where m is the number of clusters. Then, $D_{\min} = \min(D_1, \ldots, D_m)$ represents the minimum distance from the feature vector for the current waterjet encounter. $D_{\min}$ is determined for all landmines and harmless objects from the training data. Let $A = \{A_1, \ldots, A_r\}$ represent the set of minimum distances for the landmine-waterjet encounters from the training data to the nearest landmine cluster, where r is the number of landmine clusters. Let $B = \{B_1, \ldots, B_s\}$ denote the corresponding set of minimum distances for the nonlandmine waterjet training encounters. The confidence value assigned for each encounter was assigned as

$$C$$

$$= \begin{cases} 1 & \text{for } D_{\min} < B_{\min} \\ \dfrac{A_{\max} - 0.5 B_{\min} - 0.5 D_{\min}}{A_{\max} - B_{\min}} & \text{for } B_{\min} \le D_{\min} < 2 A_{\max} - B_{\min}, \\ 0 & \text{for } D_{\min} \ge 2 A_{\max} - B_{\min}, \end{cases}$$

$$(3)$$

where $A_{\max} = \max\{A_1, \ldots, A_r\}$ and $B_{\min} = \min\{B_1, \ldots, B_s\}$. C is assigned the value 1 for distances less than the minimum distance found for non-landmines (i.e., the encounter was with a harmless object) and declines linearly to 0 based on the maximum distance determined for landmines.

2.2. Maximum likelihood applied to power spectral density

The second approach investigated used the power spectral density of the sound produced by the waterjet encounter to detect landmines. This approach is a classic method used to detect and classify a signal in a noisy, indeterminate environment. It was tested because it is simple to apply and works well for a broad set of problems. Probability density functions were generated for the signal power as a function of frequency for different types of encounters. Object detection and classification was based on an ML decision.

Previous research has shown that the sampled microphone data, $r[n]$, becomes quasistationary approximately 250 ms after the waterjet is turned on over dry sand [10]. Within the quasi-stationary period, $r[n]$ can be modeled well as a Gaussian stationary random process [16]. As such, $r[n]$ can be characterized by its power spectrum, $S_r(f)$. The power spectrum derived from any particular signal will depend on a set of physical parameters, θ, such as object type, depth, and soil condition. In discrete form, the probability density function for a particular parameter set θ_i is given by

$$f(x, \theta_i) = \frac{1}{|C_i|^{1/2} (2\pi)^{k/2}} e^{-1/2 (x - \overline{x}_i)^T C_i^{-1} (x - \overline{x}_i)}, \qquad (4)$$

where

$$x = \begin{bmatrix} S_r[f_0] \\ S_r[f_1] \\ \vdots \\ S_r[f_k] \end{bmatrix} \qquad (5)$$

is a vector of measured power spectral density values at discrete frequencies f_0 through f_k, k is the number of discrete frequencies available, and $\overline{x}_i$ and C_i are the vector mean and cross correlation matrix, respectively, of the power spectral density associated with physical parameter set θ_i. For our tests, the parameters $\overline{x}_i$ and C_i were estimated from calibration data [17].

A widely accepted solution for the best choice among the set of simple hypotheses $\{H_j\}$ is given by the hypothesis, H_i, for which [17]

$$f(x, \theta_i) \geq f(x, \theta_j) \quad \forall j, \tag{6}$$

where the search space $\{\theta_j\}$ is defined over all possible physical parameters that may be encountered in a particular test. The hypothesis H_i is an "ML" solution.

Datasets used in this study were small, so principal component analysis was used to improve results. In this case [18],

$$f(x, \theta_i) = \frac{1}{\left| \Lambda_i' \right|^{1/2} (2\pi)^{j/2}} e^{-1/2(x-\overline{x}_i)^T U' \Lambda'^{-1} U'^T (x-\overline{x}_i)}, \tag{7}$$

where U is a matrix of eigenvectors, Λ is a diagonal matrix of eigenvalues, λ_i, and $\hat{C}_i = U\Lambda U^T$. The principal components of $\hat{C}_i$ are given by the eigenvalues $\lambda_0, \ldots, \lambda_j$ for which $\lambda_j > \varepsilon$, where ε is a constant chosen heuristically. The number of principal components may vary between parameter sets for a given constant ε. A change in the number of principal components causes a fundamental change in the value of the probability density function. Since the components are orthogonal, this change can be seen by the decomposition of $f(x, \theta_i)$ as the joint probability of individual components λ_j:

$$f(x, \theta_i) = \prod_j f_{\lambda_j}(x, \theta_i), \tag{8}$$

where

$$f_{\lambda_j}(x, \theta_i) = \frac{1}{\lambda_j^{1/2} (2\pi)^{1/2}} e^{-1/2(x-\overline{x}_i)^T u_j \lambda_j^{-1} u_j^T (x-\overline{x}_i)}. \tag{9}$$

Representation of one hypothesis with more principal components, j, than another places a more restrictive condition on the hypothesis with more principal components since the data must align well along more component directions. To accurately compare values of probability density between parameter sets with a different number of principal components, the jth root of the probability density function was taken before comparison. In this way, we are effectively calculating the geometric mean among values of the probability density function for each principal component and using that geometric mean to compare hypotheses.

2.3. Hidden Markov model approach

The third approach investigated was based on an HMM of the dynamics of the waterjet-soil-object interaction. The observation feature vector for discrimination is based on linear prediction coefficients and cepstral analysis which captures the local time-variant spectral characteristics of the waterjet-soil-object interaction.

The use of HMMs for object detection is motivated by the characteristics of the waterjet-soil-object interaction. Figure 1 shows a simple illustration of the waterjet setup and expected waterjet-soil-object interaction. We describe any acoustic signal as a combination of three states corresponding to the following ones:

State 1: interaction of jet with soil.
State 2: interaction of the jet with the object (when present).
State 3: decay of the jet.

The presence of the object is dictated by the presence or absence of State 2. Also, the probability of the presence of the subsequent state is dependent on the current state of the model, which is a first-order Markov model. Neither of these states are visible to the user; the user only hears the acoustic signal produced. These states show themselves as a function of the acoustic signal that is picked up by the microphone, thus the name hidden states, and hidden Markov models.

The HMM for a given object is described in terms of the probabilities of a state transition from one state to the other and the probability of the state given an observation signal [19, 20]. These probabilities and hence the HMM's can be learned using signals emitted from known objects within calibration lanes. The first step in defining the HMM is the feature selection and generation of the observation sequence.

The observation signal is the sound produced by the waterjet-soil-object interaction during the firing of the waterjet pulse. This raw acoustic signal is reduced to an observation sequence consisting of multidimensional feature vectors that capture the evolution of the waterjet-soil-object interaction. For the current research we have adopted cepstral analysis to define the feature vector for the waterjet signal that is then used by the HMM to classify that signal, though it is possible that several other feature-extraction tools may work just as well. Similar features are often used in speech processing for speech recognition and analysis [19].

Cepstral coefficients characterize the logarithm of the amplitude spectrum of the observed signal and are thus better suited for our detection problem when compared to the linear predictive coefficients themselves. The waterjet could be thought of as a source signal (impact). The recorded sound at the microphone can be thought of as the response of the buried object to this waterjet (impact) signal. The characteristic signature of this object could then be modeled in terms of its impulse response $b(t)$. Assuming that the source signal of the waterjet is $s(t)$, the recorded signal $x(t)$ is given by

$$x(t) = b(t)^* s(t) + \eta(t) \text{ or } X(f) = B(f)S(f) + N(f), \tag{10}$$

where $\eta(t)$ is an additive noise component which may be due to the background noise (such as that from the high-pressure pump) or the waterjet exiting the nozzle. For the purposes of the current discussion, we will assume that this component can either be neglected or has been filtered beforehand. Note that the spectral characteristics of the source signal $s(t)$ are not fixed and may vary due to factors such as change in waterjet pressure and variation in the standoff distance from the
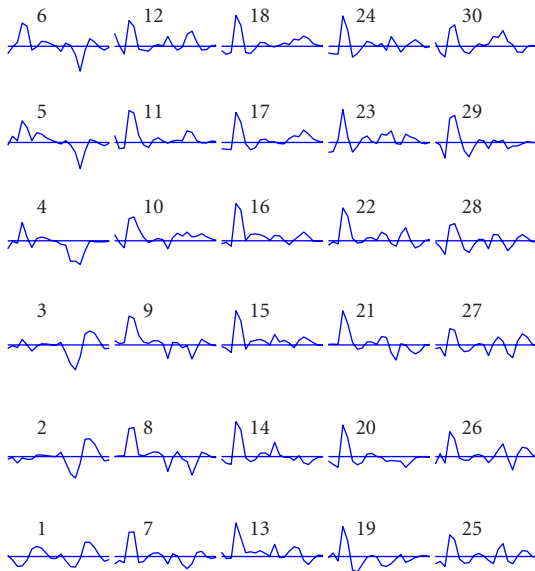
FIGURE 3: Plot showing the evolution of feature vectors with time for the signal produced by the background.

FIGURE 4: Plot showing the evolution of feature vectors with time for the signal produced by a mine (low metal antipersonnel mine).

nozzle to the surface and/or object. The quantity of interest here is the signature of the object modeled by $b(t)$ while the source signal $s(t)$ could be considered as undesirable noise which could obscure this signature. The logarithm of the amplitude spectrum of the observed signal is given by

$$\log |X(f)| \approx \log |B(f)| + \log |S(f)|. \qquad (11)$$

Thus, while variation in the spectrum of the source signal will affect the spectrum of the observed signal in a multiplicative manner, the corresponding effect on the logarithm of the spectrum is additive. As a result, the cepstral coefficients are more robust to variations in the source signal.

Figures 3 and 4 show the plot of a sequence of feature vectors for waterjet-induced signals corresponding to background-only noise and impact with the mine, respectively. Each subplot in these figures shows the feature vector $r_k = \{C_k, \Delta C_k\}$ over time for each block of the signal that is processed, where "k" is the block number ranging from 1 to T ($T = 30$), where T is the number of overlapping blocks per squirt, and C_k and ΔC_k are the cepstral and delta cepstral coefficients for the kth block, respectively. The set of all feature vectors for a given pulse define the raw observation sequence $R_n = \{r_1, r_2, \ldots, r_T\}$, where subscript n represents the nth squirt. In Figures 3 and 4, feature vectors for each block are displayed in bottom-to-top, left-to-right order. Each block is numbered for convenience.

Comparing Figures 3 and 4, we can clearly see the differences between the shape of the cepstral feature vectors associated with the background and the mine. Also note that the feature vectors are very similar for approximately the first 4 frames which show that the starting portion of the pulse for separate firings over different objects share similar characteristics. This duration may however depend on the depth of the buried object, waterjet pressure, and other factors.
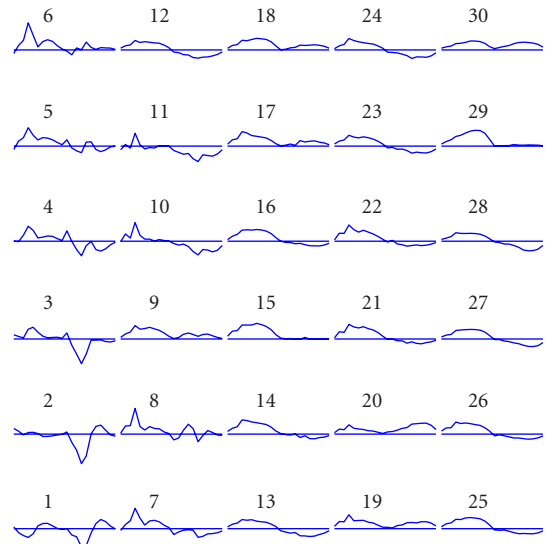
An HMM is characterized by three sets of probability matrices: the transition probability matrix (A), the observation probability matrix (B), and a prior probability matrix (Π). For the current analysis we have assumed that the system always starts in state "one" so that the prior probability matrix is fixed. Given the current state, the transition probability matrix gives the probability of occurrence of the new state. Also for a given state, the observation probability matrix assigns a probability to the occurrence of the new observation feature vector. In order to avoid computational complexity associated with continuous observation probability density functions, the feature vectors in the observation sequence are often quantized into a set of finite symbols using vector quantization. The symbols are assigned according to a minimum distance to the prototype vectors stored in a *codebook* ($\aleph$) [20]. The codebook can be estimated using the available calibration data. Given the raw observation sequence $R_n = \{r_1, r_2, \ldots, r_T\}$, the discrete observation sequence is obtained using vector quantization as $O_n = \{o_1, o_2, \ldots, o_T\}$ so that

$$o_k = VQ\{r_k, \aleph\}, \qquad o_k \in V = \{v_1, v_2, \ldots, v_M\}, \qquad (12)$$

where V is the set of all possible observation symbols and operator $VQ\{r_k, \aleph\}$ represents the vector quantization process for the given observation r_k and the codebook ($\aleph$).

An HMM for the system with N states and M observation symbols is parameterized in terms of three probability matrices A, B, and Π. We use the notation, $\Lambda = \{A_{N \times N}, B_{N \times M}, \Pi_{1 \times N}\}$ to indicate the complete parameter set of the model. Given a set of observation sequences for the system, the HMM parameter $\Lambda = \{A_{N \times N}, B_{N \times M}, \Pi_{1 \times N}\}$ can be estimated using the Baum-Walsh method [19]. In general, we would expect different Markov models for different types of buried objects (due to different characteristics of notional State 2 described earlier).

Given the HMM for class l, $\Lambda_l = \{A_{N \times N}, B_{N \times M}, \Pi_{1 \times N}\}$, the probability that the observation sequence $O_n = \{o_1, o_2, \ldots, o_T\}$ is a result of a first-order Markov process defined by Λ_l is given by the conditional probability of class l given Λ_l and O_n:

$$P(l \,|\, O_n, \Lambda_l) = P(O_n \,|\, \hat{Q}_n, \Lambda_l) P(\hat{Q}_n \,|\, \Lambda) \tag{13}$$
$$= \pi_{q_1} \prod_{k=1}^{T} b_{q_k o_k} a_{q_{k-1} q_k},$$

where π_{q_1} is the prior probability of state q_1, $b_{q_k o_k}$ is the probability of observation o_k in state q_k and $a_{q_{k-1} q_k}$ is the probability of transition from state q_{k-1} to q_k. $\hat{Q}_n$ is the optimal sequence of states $Q_n = \{q_1, q_2, \ldots, q_T\}$ that maximizes the conditional probability $P(l|O_n, \Lambda_l)$. Thus,

$$\hat{Q}_n = \arg\max_{Q_n} \left[P(l \,|\, O_n, \Lambda_l) \right], \quad Q_n = \{q_1, q_1, \ldots, q_T\}. \tag{14}$$

For waterjet-based detection purposes, an HMM is estimated for each class of object to be detected. Once the HMM has been learned for a given class or identity of object (for example, a given mine or a given class of mines), a new observation is said to belong to class l if the conditional probability $p(1|O_n, \Lambda_l)$ is above some threshold. For a multiclassification problem, the above conditional probability can be obtained for each class of objects and the class with highest conditional probability defines the identity of the buried object. Thus

$$L = \arg\max_{l} \left[P(l \,|\, O_n, \Lambda_l) \right], \quad l \in \{\text{classification}\}. \tag{15}$$

3. LABORATORY DATA

Mine detection algorithms were tested both using laboratory data and field data. Laboratory data was used to test the algorithms' ability to detect when an object was struck by the waterjet as opposed to when the waterjet struck only soil or sand. It is important to be able to distinguish a miss from a hit so the user knows when an object has been struck and because a human operator can construct a mental picture of the object's size and shape simply by striking the object several times at different locations (as is often done with a titanium probe). Such a method could also be very useful for showing if an MD has indicated a large object that is potentially a mine or a small bit of metallic debris. Field data was used to test the algorithms' ability to classify the type of object struck.

The following section details the methods and results related to the laboratory data. Field data are discussed afterwards in another section.

3.1. Methods

Laboratory data was taken from objects buried in a sand-filled tub, as illustrated in Figure 5. Objects (either a rock or dummy antipersonnel landmine) were buried approximately 1.5″ below the sand. Objects were approximately 3″ to 4″ in diameter. The waterjet was fired into the sand approximately every 2″. Location and firing of the jet was

FIGURE 5: Data was taken in the laboratory using the setup shown. Sounds produced by the waterjet-soil-object interaction were recorded by the microphone on the left. The position and firing of the waterjet nozzle (right) were controlled by a computer.

controlled automatically through a computer control system. Sounds were sampled and recorded with 16 bits of precision at 44.1 kHz using a Peavey cardioid unidirectional microphone. Water pressure was approximately 3000 psi. The waterjet was turned on for approximately 1 second for each squirt. Nozzle diameter was 0.043″. A total of 29 recordings were made of a waterjet encounter with an object and 163 of an encounter with only sand. Each recording contained a single firing of the waterjet.

For testing purposes, 10 sets of test and training data were prepared from the laboratory data. For each set, 20% of the data (20% of the object encounters and 20% of sand-only encounters) were randomly allocated for testing and 80% were allocated for training. The ability of each algorithm to detect buried objects was measured using these datasets. Results are reported for the average performance among these sets.

3.2. Results

Receiver operating characteristic (ROC) curves were calculated for each detection algorithm based on its ability to detect when the waterjet hit an object. ROC curves are given for the WDD, ML, and HMM approaches in Figure 6. The probability of false alarm necessary to reach a 90% probability of detection was 0.18 for the WDD approach, 0.25 for the maximum likelihood approach, and 0.56 for the HMM approach.

4. FIELD DATA

Field data was used to determine the ability of the algorithms to classify the type of object struck by the waterjet. Data was first taken in calibration lanes, where the type of object was known at each position. This calibration data was used to improve and train our algorithms. Data was next taken in blind test lanes, where only the approximate position of buried
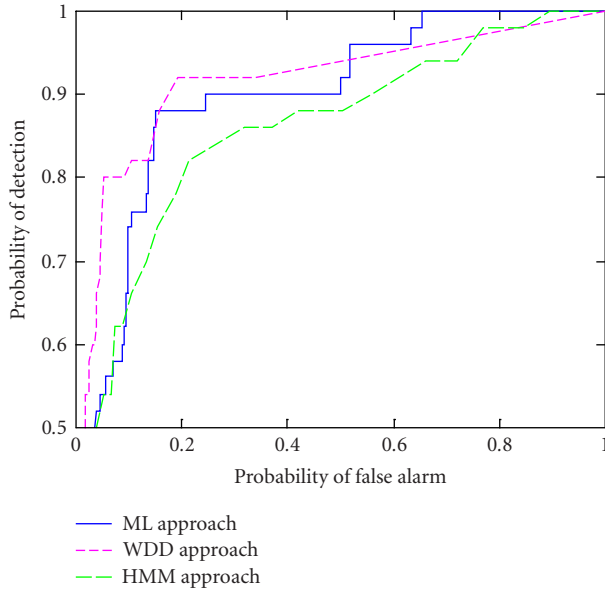
FIGURE 6: Receiver operating characteristic curve showing the ability of each approach (ML, WDD, HMM) to detect when the waterjet struck a buried object. Results are shown for data taken in the laboratory.

objects was known. Data from the blind test lanes was used to show the efficacy of the methods. The methods and results are discussed below. Because each algorithm has its own peculiar strength and weaknesses, the tests and preprocessing methods applied to the calibration data will differ from one algorithm to another.

4.1. Hardware

A hand-held "lance," shown in Figure 7, was constructed to gather field data.[1] The lance was constructed to allow an individual deminer to survey the field, giving him great freedom in the placement and number of test shots used. The lance is connected through hoses to a high-pressure pump and reservoir. A test shot is made every time the deminer presses the trigger. The length of the shot is controlled by an electronic timer and a solenoid valve mounted on the lance. Our tests used a waterjet pressure of 2000–2500 psi, a 0.05″ diameter nozzle, and squirt duration of approximately 1 second. For this setup, each squirt used approximately 2.2 cm^3 of water and penetrated the soil approximately 6″. The nozzle size and duration can be reduced to limit water usage, but even at this volume a deminer could work all day using only a few gallons of water. Sounds from each squirt were recorded by a Schoeps CCM41 supercardioid microphone mounted on the lance arm. Sounds were sampled at 96 kHz using a 24 bit digital-to-analog converter. Before each shot, the wand was placed firmly on the ground and supported by the tripod mounts. The angle between the nozzle and ground varied be-

tween 30 and 45 degrees. While this firing angle differed from the angle used in our laboratory tests shown earlier, preliminary studies in the laboratory indicate that the angle should not prevent detection and discrimination. The more shallow firing angle was required for other tests we performed using radar as part of another study.

4.2. Calibration and test lanes

Test and calibration lanes were provided for sand and for clay at a government test facility. Each lane contained 10 buried objects. Five objects were buried at a particular depth for calibration and five for test. Objects included 7 types of landmines and 3 types of harmless objects, as given in Table 1. Landmines were primarily antipersonnel-type mines, usually with very low metal content, though one antitank mine was included in the study.[2] Mines ranged in size from antipersonnel mines approximately 3″ in diameter to an anti-tank mine approximately 12″ in diameter. No specific object or mine type was repeated in a particular calibration lane. The location of each object in the lane was identified with a flag. The identity of objects next to each flag was given to UMR for the calibration sites. Test sites were constructed under the same conditions and from the same objects as calibration sites, but the object at a particular site was unknown to UMR; hence there were five "unknown" objects buried at 2″ and 4″ in both sand and clay (20 unknown objects total). Objects at "blind" test sites were identified for UMR after analysis was complete. The depths of test objects were known for clay but were unknown for sand (either 2″ or 4″ as at calibration sites).

4.3. Data

A total of 52 acoustic signals were collected from *calibration* sites for objects as well as five signals for the waterjet hitting only clay (no object clutter) and three signals for sand only (no object clutter). There were 26 waterjet-object encounters in clay and in sand each. Multiple shots were taken at each object. After squirting an object in the calibration lane, it was manually confirmed that the shot actually hit the desired object. No confirmation of a hit or miss was taken at blind *test* sites, as such confirmation could not be done during actual demining. At test sites, hits or misses were determined from the recorded sound using our algorithms. For this reason, it is possible that some recordings at test sites were classified as hitting an object when, in fact, they did not. This possibility may skew classification results shown later, but is true to what would occur during an actual demining operation.

4.4. Preprocessing and filtering of the acoustic signal

When the waterjet is fired, a low frequency vibration is induced in the wand due to the opening and closing of the waterjet valve. This vibration is picked up by the microphone due to its high sensitivity. This low-frequency vibration was

---

[1]The lance was designed by Dr. Grzegorz Galecki and Dr. David Summers of the UMR Rock Mechanics Laboratory.

[2]An agreement with our sponsor prevents us from specifying the precise mines used in the study.

FIGURE 7: The waterjet lance used to collect data in the field.

TABLE 1: Type of object located at each flag position in field calibration lanes.

Flag number	2″ sand calibration lane	4″ sand calibration lane	2″ clay calibration lane	4″ clay calibration lane
1	Antipersonnel	Antipersonnel	Metal disc	Metal disc
2	Antipersonnel	Plastic disc	Plastic disc	Plastic disc
3	Wood block	Antipersonnel	Antipersonnel	Wood block
4	Antipersonnel	Metal disc	Antipersonnel	Antipersonnel
5	Antipersonnel	Antipersonnel	Antitank	Antipersonnel

found to be additive with sounds picked up by the microphone so that we were able to filter away this contribution. A high pass 2048-tap FIR filter with a cutoff frequency of 100 Hz was used to remove this signal. Since our tests indicate there is typically no useful information in the frequency range of approximately 0–120 Hz, we were able to do this preprocessing without any loss of useful information.

4.5. Object classification

Data from calibration sites was used to train each classification approach and determine optimal processing methods. Once training was complete, the approaches were used to classify the sounds from the blind test sites. Identity of the objects at blind test sites was revealed to the authors after classification was complete. A discussion of the results of training and optimizing algorithms using calibration data follows.

4.5.1. WDD approach—calibration

Two classification approaches were investigated. First, individual models were developed for sand and for clay based on a K-means, nearest-neighbor-based discriminator. Second, a single model were developed that combined the clay and sand encounters into a single dataset. Experiments were performed to compare classification results using the two models. For the separate models, the 2″ and 4″ sand calibration data was used to train a WDD "sand + landmine" model. Likewise, the 2″ and 4″ clay calibration landmine encounters was used to train a WDD "clay + landmine" model. Soil-only encounters were used to normalize data within each soil type.

Data was normalized by subtracting the mean of the soil-only encounter for the specific soil type and dividing by the standard deviation. WDD features were computed from the normalized data. For the combined-soil-type model, the sand and clay encounters were combined to generate one dataset from which the WDD landmine model was developed. For the combined-soil type, the means and standard deviations determined from the sand-only and clay-only data were used to normalize the respective sand and clay data. For evaluation purposes, all landmine encounters were used for training. During testing, the Euclidean distance to the nearest representative landmine cluster was calculated for each encounter. Distances were used to classify objects as harmless or as landmines. ROC curves were used to evaluate results.

Experimental results showed that the combined-soil model discriminated between landmines and harmless objects better than the separate-soil models did. However, the overall landmine classification rates were poor. Setting the threshold to achieve 100% correct landmine recognition yielded 27.7% correct harmless object classification. Setting the threshold to achieve 62.0% correct landmine classification yielded 83.3% correct harmless object classification.

Experimental results for the combined soil model showed that classification rates for the first squirt at each object were much better than for the remaining squirts. Specifically, classification using the first encounter at each flag position yielded 92.3% correct landmine recognition with 72.7% correct harmless object recognition. The first shot may be a better predictor because each shot causes some changes to the soil conditions that are reflected in the sounds

TABLE 2: Percentage objects correctly identified in field calibration dataset using ML approach. In this case, the test data was taken from the same dataset used to form test statistics.

Preprocessing method	Percent correctly identified		
	Grouping 1 soil, object, depth	Grouping 2 object type	Grouping 3 mine/harmless
None	19%	37%	53%
Normalize	11%	47%	83%
Log	10%	25%	92%
Normalize and Log	12%	24%	93%

produced on subsequent firings. Accordingly, the following approach was used for classifying the blind test encounters. The combined-soil WDD feature-based landmine model was used. Test data was normalized as before. The first encounter or squirt at each flag location was used as the basis for the landmine/harmless object classification decision. The same distance thresholds were used to classify test data as with calibration data. If the Euclidean distance was less than or equal to the threshold, the encounter would be labeled as a landmine. Otherwise, the encounter was called a harmless object. If the encounter was labeled as a landmine, the type of landmine assigned to the encounter would simply be the landmine type from the calibration encounters with the closest Euclidean distance. If the encounter was labeled as a harmless object, the type of harmless object assigned to the encounter would simply be the harmless object type from the calibration encounters with the closest Euclidean distance.

4.5.2. *Maximum likelihood approach—calibration*

The maximum likelihood approach allows a grouping of data types that may be difficult to obtain with the other classification techniques. Since our calibration data was limited, the ability to form larger groups that may be independent of one or more physical parameters (for example depth or soil type) may allow the formation of better test statistics. Several possible groupings of the data were tested.

(i) Grouping 1. Data was grouped according to soil type (sand, clay), specific identity, and depth. For example, encounters with a wooden block buried at 2″ in sand would be used to generate one set of statistics. Encounters with a wooden block buried at 4″ in sand would be used to generate another. Results thus included identification of the object, depth, and soil type. In this case, objects were classified as belonging to one of 20 different groups.

(ii) Grouping 2. Data was grouped together according to object type (e.g., wood block versus plastic plate), regardless of the depth of the object or the type of soil the object was placed in. Objects were classified as belonging to one of 11 different groups.

(iii) Grouping 3. Data was grouped into two classes, landmine or harmless object.

Optimal preprocessing of data may also improve results. Three methods of preprocessing the data before application

of the ML approach were tested: (a) normalization of the power spectral density such that the integral of power spectral density evaluated to one for each measured signal, (b) taking the log of the power spectral density, and (c) first normalizing and then taking the log of the power spectral density. These techniques were also compared to the case where no preprocessing was done.

All available calibration data was used for initial training and testing. Calibration tests should still reflect performance reasonably well since data is represented statistically using only a few components and thus the approach cannot "memorize" the training set. Test signals were associated with a group according to whichever group had the highest-valued probability density function as shown in (4).

Results for the calibration dataset are shown in Table 2. The ML approach was able to correctly classify 93% of objects as harmless or harmful by normalizing and taking the log of data and was able to predict the object identity with up to a 47% accuracy by normalizing data before processing.

4.5.3. *HMM approach—calibration*

To make the estimation of LPC/cepstral coefficients less noisy and more representative of the desired signal, the original 44.1 kHz raw data were downsampled to a 6000 Hz signal. Earlier analysis has shown that the discriminatory information is predominantly in the lower frequency spectrum of the waterjet-induced acoustic signal. Up to 8th-order LPC coefficients were used for the feature vector so that the resulting feature vector was 22-dimensional.

As discussed earlier, a discrete HMM with finite observation symbols describing three states was used. A major issue in vector quantization was the design of an appropriate codebook for quantization. After some trials we found a codebook size of 64 to be appropriate for this application (i.e., there were 64 possible observations in each state). A larger codebook was not possible because we were working with a very limited dataset. Separate codebooks were designed for different soil conditions and different depths. To design the codebook, we selected an equal number of raw observation sequences corresponding to mines and harmless objects. The feature vectors for all these observations were concatenated and passed on as a representative training sequence to a program that designs the codebook using a K-means segmentation algorithm [21]. A Euclidean distance metric was used in the generation of the codebook and for code assignment.

TABLE 3: Percentage of objects correctly classified as harmful or harmless at blind field test sites.

	WDD prediction	ML prediction	HMM prediction	Observer
Sand, mixed depth	50%	50%	40%	90%
Soil, 2″ depth	60%	60%	60%	60%
Soil, 4″ depth	60%	60%	20%	60%

TABLE 4: Percentage of objects correctly identified (e.g., a wooden block or a rock) at blind field test sites.

	WDD prediction	ML prediction	HMM prediction	Observer
Sand, mixed depth	20%	10%	10%	70%
Soil, 2″ depth	20%	20%	40%	40%
Soil, 4″ depth	20%	20%	20%	20%

A separate HMM was trained for each desired classification of the targets. The calibration dataset was used to train these HMMs. The following are the steps involved in the training of the discrete HMMs.

(1) The number of states in our model was kept fixed at $N = 3$.

(2) The transition matrix and the observation matrix were randomly initialized. The a priori probabilities of the states were initialized to $\Pi = \{1, 0, 0\}$, forcing the condition that the HMM always started in State 1.

(3) All squirts corresponding to the given class were selected and the corresponding observation sequence was obtained.

(4) The quantized observation sequence was used to train the state transition matrix and observation matrix starting from the randomly initialized parameters using the Baum-Walsh method [19].

(5) Since the HMM parameter estimation may be trapped in local minima, we performed the training routine many times (with different initial conditions) and chose the model that had the maximum mean likelihood ratio.

Mine detection and classification was carried out at two levels. First, each squirt from the waterjet was classified as hitting either a mine or harmless object. Three separate HMMs were trained using calibration data for each class and each dataset. Second, after classifying the data into the classes of mine and harmless object, we proceeded to try and identify the target type (from among the seven mine types and three harmless object types) present in each data class. In this case the signals from each dataset were classified based on their target identity and separate HMMs were trained for each target type. For the soil calibration data at 2″ this results in 5 classes (4 mine types, one harmless object). Similarly for the soil calibration data at 4″ we created 5 classes and the sand calibration data generated 8 classes. After training, the HMMs were tested on the dataset on which they were trained, to check if they had been trained properly.

When testing the HMMs using the calibration training set, the HMM approach was able to correctly identify 100% of sounds as associated with a mine or harmless object and was able to correctly predict the target identity for 92% of the sounds. These results indicate that the training was accomplished effectively.

4.5.4. Blind test site results

Sounds at the blind test sites were classified using the WDD, ML, and HMM approaches as given in the previous sections. Table 3 shows the percentage of objects correctly classified as harmful or harmless for each technique. The percentage of objects whose specific identity (e.g., wooden block as opposed to rock) was correctly predicted by the algorithms is given in Table 4. These tables also include the performance of a human observer who participated in the tests and made predictions about the mine type based on what they heard or saw. The human observer did not know which object was being struck until after results had been compiled and the tests were complete.

5. DISCUSSION AND CONCLUSIONS

The goal of this study was to show the potential of using the sound produced by the impact of a high-pressure waterjet to detect and classify buried landmines. Previous work had shown this possibility existed, but did not show a clear route toward achieving accurate classification [9, 10]. In the absence of additional direction, three methods based on the temporal (WDD), spectral (ML), and a combination of temporal and spectral (HMM) characteristics were attempted. Results with laboratory data suggest the low-frequency variation of the sound signal over time is a better indication of when the waterjet hit or missed a buried object, as the WDD approach slightly outperformed the other approaches in this case. All three approaches performed similarly when attempting to classify buried objects in field experiments. The comparison in the field is a bit weak, however, due to the small quantity of data available. A clear picture of the characteristics in the sound that best identifies the buried object is still in question. The presence of these characteristics is indicated by the performance of the human observer in our tests. Finding these characteristics remains for future studies.

Classification techniques performed well when identifying whether the waterjet struck an object or hit only soil (i.e., identifying a hit/miss or object/no-object). Techniques also performed well with calibration training data when classifying encounters as with a mine or harmless object or identifying the object, but performed poorly when using data from the blind test sites. Poor performance at the blind test sites was probably related to the quantity and quality of calibration data. Each technique requires a fair amount of calibration data for appropriate training. The amount of training and test data available from our field study was relatively small. With more data, we would expect better performance.

It is interesting to note that the human observer was generally able to classify waterjet signals better than our signal processing algorithms, at least when classifying the object struck by the waterjet. Humans have an amazing ability to recognize patterns in audio signals. They also have the advantage that they may incorporate visual information into their decision, such as the location of each hit or miss when interrogating a buried object. The performance of the human observer indicates that there is additional information in the waterjet data that has not been exploited by our algorithms. Recognizing this information is a key to improving results.

The preprocessing methods and classification techniques used in these experiments were formed heuristically. Better results could be expected if techniques were based on the physics behind sound production. Sounds from the waterjet/object impact are a function of the interaction of the waterjet and object, the physical characteristics of the object, the surrounding media, the borehole created by the waterjet, and more. Understanding how sounds recorded by the microphone were produced would improve our ability to process data and extract identifying information. This understanding could be used to develop "filters" to remove unwanted information and produce measures related to the physical characteristics of the object.

In our tests, objects were classified by separately classifying the sounds from each individual squirt. The ML approach could easily be extended to make decisions based on all the squirts at an object, rather than each squirt separately. If the sounds made by a squirt at an object is independent of other squirts, then the joint probability density function for these sounds is given by

$$f(x_1, x_2, \ldots, x_k, \theta_i) = f(x_1, \theta_i) f(x_2, \theta_i) \cdots f(x_k, \theta_i), \quad (16)$$

where $f(x_1, \theta_i)$ is the probability density function for an individual squirt on object i. For the set of sounds $\{x_k\}$, the ML prediction is given by the hypothesis, H_i, for which

$$f(x_1, x_2, \ldots, x_k, \theta_i) \geq f(x_1, x_2, \ldots, x_k, \theta_j) \quad \forall j. \quad (17)$$

Using all the shots over a single object to identify the object within the calibration set improved identification of the object from 47% (single-shot classification) to 57%.

Using a waterjet to detect and classify buried objects is a unique approach to demining. Hits or misses were classified in laboratory data with more than a 90% probability of detection while achieving less than a 20% false-alarm rate. The type of object was correctly classified in up to 100% of cases when using calibration data and up to 60% of cases using blind test data. Results in the field were best with a human observer, who was able to classify objects with up to 90% accuracy. While better detection is needed for actual demining purposes, these preliminary results show the promise of the waterjet approach. Future research into the mechanisms that generate the sounds and into refinement of our classification algorithms should yield better results. The waterjet may be particularly useful as a confirmation sensor used with other sensors, like an MD. In this case, the ability to quickly and safely discern the size of a buried object or whether the object is harmful or harmless could significantly improve the demining process.

REFERENCES

[1] C. Bruschini and B. Gros, "A survey on sensor technology for landmine detection," *Journal of Humanitarian Demining*, vol. 2, no. 1, 1998.

[2] J. MacDonald, J. R. Lockwood, J. McFee, et al., *Alternatives for Landmine Detection*, RAND, Santa Monica, Calif, USA, 2003.

[3] L. Carin, Ed., "Special issue on landmine and UXO detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 6, 2001.

[4] E. Cespedes and D. Daniels, Eds., "Special issue on UXO and mine detection," *Subsurface Sensing Technologies and Applications*, vol. 2, no. 3, 2001.

[5] K. Bruschini, C. De Bruyn, H. Sahli, and J. Cornelis, "EUDEM: The EU in humanitarian DEMining—Final Report," 1999, http://www.eudem.vub.ac.be/publications/.

[6] N. Milisavljevic, "Comparison of three methods for shape recognition in the case of mine detection," *Pattern Recognition Letters*, vol. 20, no. 11–13, pp. 1079–1083, 1999.

[7] N. Milisavljevic, I. Bloch, and M. Acheroy, "Modeling, combining, and discounting mine detection sensors within the Dempster-Shafer framework," in *Detection and Remediation Technologies for Mines and Minelike Targets V*, vol. 4038 of *Proceedings of SPIE*, pp. 1461–1472, Orlando, Fla, USA, April 2000.

[8] N. Milisavljevic, I. Bloch, and M. Acheroy, "Characterization of mine detection sensors in terms of belief functions and their fusion, first results," in *Proc. IEEE 3rd International Conference on Information Fusion (FUSION '00)*, vol. 2, pp. THC3/15–THC3/22, Paris, France, July 2000.

[9] R. Denier, "A heuristic approach to landmine detection using pulsed waterjet excitation," M.S. thesis, University of Missouri-Rolla, Rolla, Mo, USA, 1999.

[10] J. A. Stuller, S. J. Qiu, and K. Das, "Signal processing for land mine detection using a waterjet," in *Detection and Remediation Technologies for Mines and Minelike Targets IV*, vol. 3710 of *Proceedings of SPIE*, pp. 1330–1342, Orlando, Fla, USA, 1999.

[11] D. A. Summers, *Waterjetting Technology*, E & FN Spon, London, UK, 1995.

[12] R. J. Stanley, P. D. Gader, and K. C. Ho, "Feature and decision level sensor fusion of electromagnetic induction and ground penetrating radar sensors for landmine detection with handheld units," *Information Fusion*, vol. 3, no. 3, pp. 215–223, 2002.

[13] R. J. Stanley, S. Somanchi, and P. D. Gader, "Impact of weighted density distribution function features on land mine detection using hand-held units," in *Detection and Remediation Technologies for Mines and Minelike Targets VII*, vol. 4742 of *Proceedings of SPIE*, pp. 892–902, Orlando, Fla, USA, April 2002.

[14] R. J. Stanley, N. Theera-Umpon, P. D. Gader, S. Somanchi, and D. K. Ho, "Detecting landmines using weighted density distribution function features," in *Signal Processing, Sensor Fusion, and Target Recognition X*, vol. 4380 of *Proceedings of SPIE*, pp. 135–141, Orlando, Fla, USA, April 2001.

[15] J. M. Zurada, *Introduction to Artificial Neural Systems*, West Publishing, Boston, Mass, USA, 1992.

[16] S. J. Qiu, "Acoustic landmine detection," M.S. thesis, University of Missouri-Rolla, Rolla, Mo, USA, 1999.

[17] V. K. Madisetti and D. B. Williams, Eds., *The Digital Signal Processing Handbook*, CRC Press, Boca Raton, Fla, USA, 1997.

[18] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*, Addison-Wesley, New York, NY, USA, 1990.

[19] L. R. Rabiner, "A tutorial on Hidden Markov Models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.

[20] L. R. Rabiner and R. W. Shafer, *Digital Processing of Speech Signals*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1978.

[21] Y. Linde, A. Buzo, and R. Gray, "An algorithm for vector quantizer design," *IEEE Trans. Communications*, vol. 28, no. 1, pp. 84–95, 1980.

Daryl G. Beetner is an Associate Professor of electrical and computer engineering at the University of Missouri, Rolla. He received his B.S. degree in electrical engineering from Southern Illinois University at Edwardsville in 1990. He received an M.S. and Doctor of Science degree in electrical engineering from Washington University in St. Louis in 1994 and 1997, respectively. He conducts research on a wide range of topics including electrocardiology, skin cancer detection, humanitarian demining, and electromagnetic compatibility.

R. Joe Stanley received the B.S.E.E. and M.S.E.E. degrees in electrical engineering and a Ph.D. degree in computer engineering and computer science from the University of Missouri-Columbia. As a graduate student at the University of Missouri-Columbia, he worked under training grants from the National Library of Medicine and the National Cancer Institute. Upon completing his doctoral study, he served as Principal Investigator for the Image Recognition Program at Systems & Electronics, Inc. in St. Louis, Mo. He is currently an Assistant Professor in the Department of Electrical and Computer Engineering at the University of Missouri-Rolla. His research interests include signal and image processing, pattern recognition and automation.

Sanjeev Agarwal completed his Ph.D. in Electrical and Computer Engineering from University of Missouri-Rolla in 1998 and BTech and MTech degrees (1993) from Indian Institute of Technology, Bombay. Dr. Agarwal joined the Department of Electrical and Computer Engineering at University of Missouri-Rolla in 1998 and is now a Research Assistant Professor there. Dr. Agarwal's research interests include machine vision, intelligent computing, image processing, multisensor fusion, automatic detection theory, airborne terrain analysis and reconnaissance, and virtual and augmented reality.

Deepak R. Somasundaram received his B.S. in electronics and communications engineering from University of Madras, India in May 2001. He joined the University of Missouri-Rolla in fall 2001 as a graduate student of electrical engineering, where he worked with Dr. Sanjeev Agarwal on automatic target detection using hidden Markov models. Currently Deepak is a DSP Engineer with Phonic Ear. His current research includes adaptive acoustic feedback cancellation and DSP hardware implementations.

Kopal Nema obtained her B.S. degree in electronics and telecommunications from Pune University, India and received an M.S. degree in computer engineering from University of Missouri-Rolla in May, 2003. She has previously worked as a software engineer for Cisco. She is currently working with Intel in Bangalore, India.

Bhargav Mantha received the B.S. degree in electrical and electronics engineering from Chaitanya Bharati Institute of Technology, Osmania University, in June 2001 and his M.S. degree in electrical engineering from University of Missouri-Rolla in August, 2003. He now works with Credit Suisse First Boston in New York, USA.