Camp Butner Live-Site UXO Classification Using Hierarchical Clustering and Gaussian Mixture Modeling

Alex Bijamov, Juan Pablo Fernández, Benjamin E. Barrowes, Member, IEEE, Irma Shamatava, Kevin O’Neill, and Fridon Shubitidze, Senior Member, IEEE

Abstract—We demonstrate in detail a semisupervised scheme to classify unexploded ordnance (UXO) by using as an example the data collected with a time-domain electromagnetic towed array detection system during a live-site blind test conducted at the former Camp Butner in North Carolina, USA. The model that we use to characterize targets and generate discrimination features relies on a solution of the inverse UXO problem using the orthonormalized volume magnetic source model. Unlike other classification techniques, which often rely on library matching or expert knowledge, our combined clustering/Gaussian-mixture-model approach first uses the inherent properties of the data in feature space to build a custom training list that is then used to score all unknown targets by assigning them a likelihood of being UXO. The ground truth for the most likely candidates is then requested and used to correct the model parameters and reassign the scores. The process is repeated several times until the desired statistical margin is reached, at which point a final dig is produced. Our method could decrease intervention by human experts and, as the results of the blind test show, identify all targets of interest correctly while minimizing false-alarm counts.

Index Terms—Agglomerative hierarchical clustering, Camp Butner, classification, electromagnetic induction (EMI), ESTCP, inverse problems, ONVMS, semisupervised learning, unexploded ordnance (UXO).

I. INTRODUCTION

Unexploded ordnance (UXO) are a deadly remnant of past wars and military training. Unexploded bombs and mortar rounds, dating back to World War I or even further, continue to be found and keep posing a significant humanitarian risk in our time. The character and pervasiveness of modern warfare have caused the number of UXO to grow significantly. In the U.S. alone, there are approximately ten million acres of land that were once used by the Department of Defense or the Department of Energy for training purposes and are now potentially contaminated with UXO. The detection and disposal of UXO is an expensive and time-consuming process that makes the cleaning up of contaminated land a major environmental challenge. Because it is not yet practical to detect the explosive within more or less intact buried ordnance, one must instead detect the ordnance itself, i.e., its metal structure. While it is relatively easy to detect buried metallic objects of any kind, it is not trivial to differentiate them by type: Depending on their location, potentially dangerous targets of interest (TOI) and innocuous clutter can produce magnetic-field anomalies of comparable magnitude. Typically, a more sophisticated analysis must be carried out to determine the exact nature of the source producing a given measured response.

Recent years have seen the development of various sensing technologies based on electromagnetic induction (EMI) [1]–[3], which involves a range of frequencies between tens of hertz and hundreds of kilohertz in which the ground is essentially transparent. In land-based UXO detection scenarios, one or several EMI transmitters are used to send a “primary” magnetic field through the ground, which induces eddy currents and magnetic response in nearby metallic bodies. In time-domain EMI systems, in accordance with Faraday’s law, turning the excitation off causes underground objects to produce a time-decaying (“secondary”) EMI field that is then detected by one or more receivers and analyzed. Physical differences in the shape and composition of targets (their volume, metal thickness, etc.) cause these secondary signals to have different time-decay characteristics (e.g., initial intensity, early- or late-time decay rate) even after the effects of the targets’ locations and orientations with respect to the sensor have been factored out.

The Environmental Security Technology Certification Program (ESTCP) recently launched a series of blind tests at increasingly challenging and complex live sites [4]–[6] to demonstrate the state of the art in EMI detection technologies and UXO discrimination and classification algorithms. The first test was conducted in 2007 at the former Camp Sibert in Alabama, USA, using first-generation sensors (the commercially available EM61-MK2 and EM63, both developed by Geonics Ltd.). The Sibert test was relatively simple: one had to discriminate well-isolated large intact 4.2” mortars from...
smaller range scrap, shrapnel, and cultural debris. The second ESTCP discrimination study was set up in 2009 at the UXO site in San Luis Obispo (SLO), CA, and featured both a more challenging topography and a wider mix of TOI [4], [5]. Magnetometers and first-generation EMI sensors were deployed to the site and used for an initial exploratory survey. Three advanced EMI sensing systems were then used to perform a close interrogation of the detected anomalies. Among the munitions buried at SLO were 60-mm and 81-mm projectiles, 4.2” mortars, and 2.36” rockets; three additional types of ordnance were discovered during the course of the demonstration.

In this paper, we concentrate on the third ESTCP discrimination test, which took place at the former Camp Butner, located approximately 15 mi north of the city of Durham and straddling Durham, Granville, and Person Counties, all in North Carolina [7]–[9]. The War Department acquired the 40,384-acre property from private landowners in 1942 and used it during World War II as a training and cantonment facility for infantry divisions (including the 78th, 89th, and 4th) and miscellaneous artillery and engineering units [5]. A large variety of munitions have been reported as used at the camp, including rifle grenades, 2.36” rockets, 37- and 40-mm rounds, 81-mm mortars, and 105-, 155-, and 240-mm projectiles [5]. The particular range chosen for the test was known to contain mainly M48 fuzes, 37-mm rounds (some with a copper driving band and some without), and 105-mm projectiles, along with explosion byproducts such as partial mortars (i.e., stretched-out half-shells), smaller shrapnel, and artificial metallic clutter. There were 2291 anomalies in total, of which 171 were TOI.

Although we briefly discuss the processes of data acquisition and inversion, our main objective here is to report on the performance of an automated classification approach that reduces human involvement (and human error) and attempts to process data with high sensitivity (i.e., detecting all potentially dangerous targets) and specificity (i.e., leaving buried as many innocuous targets as possible). Several clustering and statistical signal processing techniques [10] have been used previously to identify UXO starting from magnetometer and EMI data [11]–[14]. In at least one case, the classification is based on “shape and size information” extracted from direct magnetic field observations “examined visually” and “viewed as [...] image[s],” with no underlying physical model [15]. (The method presented in that reference is found to outperform a physics-based model, although that may be because the physics in the model is not quite correct.) Library matching and support vector machines or neural-net-based decision-making [16]–[23] have been studied and applied to UXO classification based on different features, such as size or temporal decay [9]. Instances of successful application of Bayesian data fusion, multivariate Gaussian representation of EMI signals, and semisupervised learning techniques have been also reported [24]–[27]. Most of these techniques, which are based on the simple dipole model, require human knowledge and intervention and need a sample of training data, which in some cases may not be representative, or even available. The work in [10] describes a procedure that looks simultaneously for relevant features and classification protocols and is thus similar in spirit and more general than the one proposed here. On the other hand, that procedure is designed and honed for tasks with high dimensionality and possibly many irrelevant features; as we shall see in the following, the data at our disposal has very few dimensions, and every feature is germane.

The goal of our automated classification process, which uses a more detailed model to invert the data and provide classification features, is to reduce the workload of human experts by having them perform the crucial tasks of decision-making and quality control while delegating to the software routine tasks such as feature extraction, clustering, and labeling. Another feature of our method is that it does not require any precollected training data. Instead, the algorithm “trains itself” by asking for a set of anomalies to be unearthed—either because their features are particularly representative of an already identified category of targets, or because they belong to a yet-unexplored region of the feature space—and then taking in the new knowledge. At the first stage of the process, the method examines the data and uses only unsupervised learning techniques to build and request a custom training set required to identify the “suspicious” regions characterizing all the types of UXO expected at the site. Once this partial truth is known, the model incorporates it and starts an iterative process, internally updating its settings to accommodate newly available information and then either requesting more ground truth to further fine-tune the classification or, at the end, providing a final priority-weighted dig list.

The third and most important feature of the method is its performance. On the Camp Butner test, our procedure yielded 100% accuracy in TOI detection and by-caliber classification; it required 295 false alarms to be dug out, thus leaving in the ground more than 85% of the clutter.

II. METHODS

A. Data Acquisition

The data presented here were acquired at Camp Butner using the time-domain electromagnetic towed array detection system (TEMTADS), an outstanding next-generation EMI sensor developed by the Naval Research Laboratory, SAIC, and G&G Sciences that had previously been used at SLO. TEMTADS consists of 25 horizontal collocated transmit/receive pairs, each composed of a 35-cm square transmitter coil surrounding a 25-cm square receiver coil, arranged in a rectangular 5 × 5 grid with 40-cm neighbor-to-neighbor separation [28] (see Fig. 1); this rigid arrangement greatly enhances the positional accuracy of the apparatus. During a measurement, the transmitters are activated one by one, in sequence, with all 25 receivers receiving; 625 transients are obtained as a result, sampled logarithmically over 123 time gates ranging from approximately 100 μs to 25 ms. The sensor thus gathers the secondary field under different illumination conditions and provides a wide and diverse array of data for processing and inversion. It was possible to collect data with TEMTADS because the site topography was rather benign, i.e., flat, grassy, and treeless.

B. Data Inversion

UXO discrimination demands a fast and accurate representation of a target’s EMI response. The model used most
frequently for this purpose approximates the whole object with a set of orthogonal colocated point dipoles whose induced effective moment is related to the primary field through a symmetric polarizability tensor [29]. The model is fast and easy to implement and interpret, but it rests on assumptions that often limit its usefulness [30]. Large and complex targets (e.g., those containing sections with different materials, fins, rings, etc.) require more advanced methods such as the normalized surface magnetic source (NSMS) model [31]–[34] that we used at Camp Sibert and SLO [23]. The NSMS method spreads a nonuniform distribution of normally oriented dipole moment (scaled by the primary field) over a virtual prolate spheroidal surface that encloses the target. This distribution can be determined directly by minimizing the difference between measured and modeled data for a known object-sensor combination at a given relative location and orientation. The integral of the distribution over the spheroidal surface is a global effective magnetic moment. To avoid these issues, the ONVMS method uses a generalization of the Gram–Schmidt procedure [37] to construct the orthonormalized sets of Green functions, it takes into account mutual couplings between different sections of the different targets and, at the same time, avoids the appearance of singular matrices in multitarget situations. It is indifferent to the number of targets: Once the amplitudes and the locations of the source dipoles are determined, one needs only to look at their spatial patterns and compute and diagonalize the time-dependent total polarizability tensor for each spatial group. The resulting time-dependent diagonal elements have a phenomenology similar to those of dipole polarizabilities [42] or total NSMS amplitudes [43], providing insight into intrinsic

\[
\mathbf{H}(\mathbf{r}) = \sum_{i=1}^{M} G_i(\mathbf{r})\mathbf{m}_i = [G_1 \ G_2 \ \ldots \ \mathbf{m}_1 \ \mathbf{m}_2 \ \ldots], \quad (2)
\]


where the total field is the superposition

\[
\mathbf{H}(\mathbf{r}) = \sum_{i=1}^{M} G_i(\mathbf{r})\mathbf{m}_i,
\]

which can be inverted via a least-squares search to find the locations and dipole moments of the sources given an array of measured \(\mathbf{H}\)-values corresponding to multiple transmitters, multiple receivers, and different vector components. This multitarget dipole model works well for one or two sources, but for larger numbers (and thus matrix sizes) becomes time-consuming and increasingly ill-posed, often requiring regularization. To avoid these issues, the ONVMS method uses a generalization of the Gram–Schmidt procedure [37] to construct a set of orthonormal Green functions \(\Psi_k\) starting from \(G_i\) and synthesizes a modeled field using

\[
\mathbf{H}(\mathbf{r}) = \sum_{k=1}^{M} \Psi_k \mathbf{b}_k,
\]

where the source-amplitude coefficients \(\mathbf{b}_k\) can be determined from measured data by inverting matrices of size \(6 \times 6\) [38]; the dipole moments \(\mathbf{m}_i\) can be then synthesized in terms of \(\mathbf{b}_k\).

During actual inversion, the model is combined with differential evolution optimization [39], [40] to iterate through various positions of the sources until a suitable discrepancy is attained between modeled and measured field [38]. Knowledge of the final locations and dipole moments allows calculation of the time-dependent polarizability tensor \(\mathbf{M}_i\) for each target [30], [35]. Finally, by combining the data from all time channels and determining their joint eigenbasis [41], we obtain temporal decay curves of the primary components of the magnetic polarizability tensors in the targets’ principal body axes: \(Q_{x,i}(t)\), \(Q_{y,i}(t)\), and \(Q_{z,i}(t)\).

The advantage of the ONVMS technique is that, by constructing the orthonormalized sets of Green functions, it takes into account mutual couplings between different sections of the different targets and, at the same time, avoids the appearance of singular matrices in multitarget situations. It is indifferent to the number of targets: Once the amplitudes and the locations of the source dipoles are determined, one needs only to look at their spatial patterns and compute and diagonalize the time-dependent total polarizability tensor for each spatial group.

Fig. 1. Schematic of the TEMTADS sensor array and its transmit/receive pairs.
target properties such as cylindrical symmetry, surface area, metal content, and ferromagnetic behavior.

The ONVMS method takes as input the number $M$ of responding sources, which is, in principle, unknown. For the MetalMapper [28] and the $2 \times 2$ 3-D and $5 \times 5$ TEMTADS arrays [28], [38], we have developed a procedure based on joint diagonalization that estimates $M$ with no need to invert any data [41]. For other sensors, one may proceed by letting $M$ vary as part of an optimization routine.

In this paper, we used one-, two-, and three-target inversions in sequence to gradually improve the accuracy of classification. One-target inversion is equivalent to the simple dipole model and can be used to perform an initial quick screening of the site and identify easy targets. Its performance degrades significantly, however, in situations where multiple physical objects fall in the field of view of the sensor. While inversions assuming more than one source can effectively decouple the signals arising from these complicated scenarios, additional processing is required when the actual number of buried objects is less than their assumed number $M$; in those cases, two or more triplets $Q_{(x,y,z)}$ of effective polarizability decay curves could correspond to different parts of a single physical object. In order to achieve consistency in object characterization and to guarantee that all dangerous targets are identified, it is important to try out all possible combinations of these curves. For $M$-target inversion, it is necessary to consider a total of $\Sigma_{k=1}^{M} \binom{M}{k} = 2^M - 1$ combinations of effective polarizability decay curves: those of individual targets $Q_{(x,y,z)}(t) = JD(\bar{M}_i(t))$, their duplets $Q_{(x,y,z)}(t) = JD(\bar{M}_i(t) + \bar{M}_j(t))$, triplets $Q_{(x,y,z)}(t) = JD(\bar{M}_i(t) + \bar{M}_j(t) + \bar{M}_k(t))$, and so on, where $i \neq j \neq k \in \{1, \ldots, M\}$ and $JD$ stands for joint diagonalization.

We observe that multitarget inversions provide target signatures that are more consistent case to case than those resulting from single-target inversions. For example, Fig. 2 shows the $Q_z$ EMI decay curves obtained for two of the Camp Butner anomalies, #30 at left and #3 at right, both of which contain 105-mm high-explosive antitank (HEAT) rounds. In each case, the black crosses denote the effective longitudinal polarizabilities obtained from one-target inversion, whereas the red dots (seven curves per plot) show the possible combinations that result from three-source inversion. For discussion purposes, let us label those sources A, B, and C. The UXO at left was almost vertical (the ground truth, as shown in the photograph, reveals that its inclination was 81°) and the one-target inversion captures only its top region, which is much closer to the sensor than the rest. On the other hand, if we allow two or more dipoles to describe the target, we see a decay curve that agrees better with the case at right and with HEAT-round decay profiles known from previous experiments. Although all three sources contribute to the signal, one of them (e.g., C) is clearly dominant; the bottom curves thus correspond to the possibilities A, B, and A + B, whereas the top curves represent C, A + C, B + C, and A + B + C. In the case at the right, the target, whose inclination is now −10°, is in full view of TEMTADS and can be described well using a single source. The three-source inversion favors a two-dipole description of the TOI in this instance: If A and B are the relevant sources, then the curves at bottom correspond to A, B, C + A, and B + C, and those on top represent A + B and A + B + C. Only by considering the combinations of sources can we expect to characterize TOI unambiguously enough to identify and classify them, particularly if they are large, complicated in structure, or shallow [44].

Let us illustrate the validity of the previous statement with another example. Consider two permeable cylinders placed edge to edge and illuminated by the field of a faraway square loop. We calculate the EM response in the frequency domain using the method of auxiliary sources [45] and translate it into the time domain using Anderson’s logarithmic filters [46]–[48]. The calculations are performed with and without incorporating the interaction between the cylinders. (In the latter case, we simply solve for each individual cylinder and add the results.) Fig. 3 shows that the interaction between the cylinders is quite noticeable, particularly at late times. Elsewhere [49], we have published a more detailed study of the effect of target–target interactions in EMI sensing; in that paper, we establish how the shapes of the targets, the distance between them, and the orientation of the primary field affect their mutual coupling.

Going back to the 105-mm HEAT rounds in Fig. 2, we further compare the true sensor readings to those reconstructed using one-source and three-source inversion. Fig. 4 shows the sensor readings from all 25 receivers corresponding to the primary...
Fig. 3. Time-domain responses from a system of two cylinders with and without incorporating their mutual interaction. The cylinders have conductivity $4 \times 10^6 \text{S/m}$, relative permeability 150, diameter $2a = 10 \text{cm}$ and length $L = 20 \text{cm}$. Their edge-to-edge separation is (left) $D = a/100$ and (right) $D = a/2$. The primary field is generated by a $1 \text{m} \times 1 \text{m}$ rectangular loop placed 1 m away from the targets. The curves labeled “No Interaction” show the sum of the individual target responses; those labeled “Interaction” display the solution of the full problem.

Fig. 4. Reconstruction of the magnetic field in all 25 TEMTADS receivers using one-source and three-source inversion for the targets of Fig. 2. The primary field is that of the central transmitter (No. 12 in Fig. 1). The three-source inversion approximates the original measured data better than the one-source model. The loss of accuracy is particularly noticeable for the case in which the target is vertical and the one-source inversion forces the dipole to shift toward the part of the target closer to the sensor.

In this paper, we have concentrated on $Q_z$, the largest effective polarizability, to perform classification; we have noticed that the other amplitudes are often compromised due to low SNR, particularly for small or deep targets [23]. (Henceforth, we suppress the subscript $z$.) We decrease the dimensionality of the feature space by fitting the decay curves to the following empirical law popularized by Pasion and Oldenburg [16]:

$$Q(t) = k t^{-b} e^{-gt}.$$  (3)

To decrease the effects of noise, we first linearize (3) by taking its logarithm. We then extract the Pasion–Oldenburg parameters $b$, $g$, and $\log k$ using nonlinear least squares, imposing the constraints $b > 0$ and $g > 0$ to preserve the physical essence of the data. The fits are always performed for times $t > 0.1 \text{ms}$ to avoid perturbations due to the primary field, whose shutoff is not instantaneous. In order to disregard the “leveling off” of the decay curves in log space when they reach the noise threshold, we also assign an upper bound to the time channels used for the fit. This bound depends on the particular anomaly, and within each anomaly differs for the $x$-, $y$-, and $z$-decay curves. To find the bound, we estimate the inherent noise level of the decay curve $Q(t)$ as its mean value over the last $N$ time channels (with $N = 10$ chosen by hand) and set the threshold value $t_{\text{max}}$ to the time in which the curve falls below this average: $Q(t_{\text{max}}) \leq \text{mean}(Q_{123}, \ldots, Q_{121})$. Due to the presence of several highly noisy anomalies, we must impose a lower bound on $t_{\text{max}}$ to ensure that there is sufficient information for the fits; we chose $t_{\text{max}} > 0.3 \text{ ms}$. These considerations result in dramatically improved fits, as shown in Fig. 5 [23].

C. Classification

In this paper, we combine agglomerative hierarchical clustering [50]–[52] and probabilistic classification to perform semiunsupervised learning for UXO discrimination. The classification features that we use are the Pasion–Oldenburg parameters extracted from ONVMS decay curves. We first used clustering in feature space to split the entire data set into a finite number of clusters. The number of clusters is an external parameter and is set to be between 1% and 5% of the number of items in the data set. There are several options for clustering based on different criteria and distance metrics; we found
the following two combinations particularly useful for initial screening of the feature space [23].

1) Ward linkage with Euclidean distances [53]–[56]. The Ward technique encourages the formation of homogeneous clusters by controlling the sum over all clusters of the sum of squared Euclidean distances between the members of a cluster and its centroid; i.e.,

$$E = \sum_{k=1}^{K} \sum_{x_j \in C_k} \|x_j - m_k\|^2,$$

where $K$ is the total number of clusters and $m_k$ is the centroid of cluster $C_k$ [55].

At each step of the process, the two clusters whose merger increases $E$ by the least amount are merged, decreasing the number of clusters by one unit. The process starts with the individual data points as clusters and stops when the desired number of clusters, set a priori to about 5% of the total number of anomalies, is reached.

2) Weighted pair group method average (WPGMA) linkage with Mahalanobis distances. Like Ward’s, this method starts with the individual data points and merges them into clusters, decreasing the number of entities one by one until a preset value is reached. The clusters joined at each step are those found to be separated by the shortest WPGMA metric, which is defined recursively [54], [55] by the rule

$$d(R, P + Q) = (d(R, P) + d(R, Q))/2$$

expressing the separation between cluster $R$ and another resulting from the merger of $P$ and $Q$ [57].

To compute the separation between data points at the last level of recursion, we use the Mahalanobis distance [58], [59], which takes into account the natural variation and spread of physically different feature values in their own dimension and is thus potentially useful for clustering.)

Upon finding the clusters, we requested the ground truth for the anomalies that lay closest to the geometric center of each. The clusters that happened to be centered around TOI were further labeled as potential UXO clusters and used as a basis to construct a Gaussian mixture model (GMM) in which each suspicious cluster was fit with a multivariate normal distribution. Every anomaly was assigned a score—a measure of its probability of being a particular type of UXO—based on its position relative to the UXO clusters; those scores, in turn, were used to sort the anomalies and generate a prioritized dig list.

As mentioned above, multitarget inversion often provides better target localization and more precise and consistent—and thus more reliable—classification features. In a two-target inversion case, however, this means having three times as many data points in feature space, because each physical anomaly is represented by a triplet of points. If any of these three points is suspected to be a UXO, the whole anomaly (i.e., all objects buried at that specific location at the UXO site) will be treated as UXO.

Although it is still possible to perform the data clustering in multitarget inversion feature space, there is no straightforward way to identify which group of points to request as training data and how to interpret the results. Suppose a training data point is requested from a certain cluster that contains only clutter (perhaps even having outlier values as features). Whereas the data point corresponding to this object indeed has features peculiar to clutter, it can happen that this point belongs to a triplet containing a UXO, which will be revealed after the ground truth for the particular anomaly is studied. Since it is impossible to determine, without prior library knowledge, which of the points in the two-target-inversion triplet (target 1, target 2, or their superposition) actually represents the signal from the UXO, one would have to flag all three locations in feature space as potential UXO and contend with an intolerably large false-alarm rate.

To get around this issue, we employ a two-step approach. In the first step, the features extracted only from single-target inversions are clustered, and the corresponding ground truth is requested. The central elements belonging to each of the clusters are then probed, and those identified as UXO are marked. The clusters associated with them are used to construct the first generation of GMMs, which are then imported into the multi-target inverted feature space. Items in the latter space are then rated as to their likelihood of being UXO, based on the initial GMMs. Ground truth is then requested for the items rated most likely to be UXO. The multivariate GMM can be then constructed around the identified UXO clusters, and the rest of the anomalies are assigned a score that quantifies their likelihood of being UXO.

The combined clustering/GMM approach therefore provides a natural way to find intrinsic patterns in noisy feature data and yields a convenient probabilistic measure of class membership for unknown items. It also reduces the amount of required training data and improves both classification sensitivity and specificity.

III. Results

In this section, we apply the classification technique described earlier to the Camp Butner blind-test data. The classification process, summarized in Fig. 6, consisted of the following steps.

1) No initial training data were requested. (We knew, however, that three types of UXO targets were expected: 37- and 105-mm projectiles and M48 fuzes.)

2) The features $b$, $g$, and $\log k$ were extracted from the data for all anomalies starting from one-target, two-target, and three-target inversion.

3) An initial Ward clustering (with Euclidean distances) was performed, and, in order to probe the feature
space, the ground truth was requested for the 69 targets whose features were located closest to the cluster centroids.

4) Clusters containing at least one UXO were identified, and a smaller domain was selected within the feature space for further interrogation. (One projectile of each kind was identified, but no fuzes.)

5) A second clustering (WPGMA, with Mahalanobis distances) was performed within the selected domain, primarily to seek for M48 fuzes. The targets with features closest to the corresponding cluster centroids (26 at this step) were singled out and their ground truth requested. The clusters with at least one identified UXO were marked as suspicious.

6) All targets whose features (based on two-object inversion) fell inside any of the suspicious clusters were used to train a three-component GMM classifier and score all of the unknown targets.

7) All targets with a score greater than a specifically selected threshold were assumed to be UXO, and the ground truth was requested for them. (Selecting this threshold is one of the few parts of the process that involves human decision-making.) This set consisted of 131 targets, three of which had already been requested previously; of these, 118 were confirmed to be UXO.

8) A new three-component GMM classifier was trained using features from the three-object EMI inversions. All the items were rescored to correct for changes and incorporate new information. Another 20 targets with scores decreasing from a certain expert-defined low value were selected for additional verification. At this point, if the verification had yielded that all the chosen targets were clutter, the algorithm would stop, and the scored values would be used to produce a final dig list.

9) Four out of the 20 items requested happened to be UXO, and the classification continued. (The ground truth for three of these 20 items had already been requested in the previous steps, and one had been confirmed as UXO.)

10) All confirmed UXO were separated into three groups (105-mm, xxx, and 37-mm); no further within-group discrimination was performed. Each of the three groups was used to train a separate one-component GMM classifier that scored all the targets with respect to target type (based on the features obtained from precise three-object EMI inversions). The ground truth was then selected individually for each object type, which is based on a certain threshold score. This step helped describe the individual clusters in more detail.

11) A total of 36 items were requested from a 105-mm scored data set, with 18 being already known; 174 items were requested from a 37-mm scored data set, with 118 being already known; and 53 items were requested from a M48 scored data set, with 27 of them being already known. At this stage, a total of 322 items were requested, of which 162 were UXO.

12) Finally, a three-component GMM classifier was trained on the confirmed UXO and further used to score all of the unknown targets. A specific threshold was then selected, and the final dig list was produced.

Fig. 7 shows the results of the first two clustering processes, corresponding to Steps 2–5 given earlier. Only four UXO targets were identified at this stage: two 37-mm munitions (with features very close to each other in Fig. 7), one M48 fuze.
Fig. 7. Clustering of the EMI features from a single-object inversion for Camp Butner. (Left) Results of the first clustering using weighted linkage with Mahalanobis distances. Dots are color-coded by cluster membership. (Right) All four identified UXO (black markers) after the second clustering round within a smaller domain ($\log k \in [2, 8], b \in [0.05, 2], g \in [0.05, 2]$) using Ward linkage and Euclidean distances. Two of the four dots belong to the same cluster of 37-mm targets.

Fig. 8. Clusters used to train the first GMM classifier and its scoring results. (Left) Assumed UXO clusters (black dots) used to generate the three-component GMM classifier. There are three clusters, one for each type of UXO. Each cluster is centered close to a target previously identified as UXO (shown as a red dot and visible only for the cluster at center) and contains other anomalies that presumably are dangerous. (Right) Histogram of the anomalies sorted by an arbitrary score that measures their likelihood of being UXO. The ground truth was requested based on thresholding the score at an externally selected value close to 0.5.

Fig. 9. Updated GMM classifier after confirming 118 UXO items in the Camp Butner data. (Left) Score isosurfaces based on the GMM classifier trained on all currently identified UXO, in the feature space corresponding to three-object EMI inversion. (Right) Score histogram using the new classifier. An additional 20 items were requested to investigate the region corresponding to $\log \text{(score)}$ within $[-6, -5]$. This region was identified by the expert after observing the corresponding score isosurfaces and the histogram.

and one 105-mm projectile. Fig. 8 illustrates the training data used to create a three-component GMM classifier in Step 6 and the resulting score distribution histogram. The external interaction from the expert in this case consisted in selecting an anomaly scoring threshold beyond which the ground truth would be requested. We picked a value of $\log \text{(score)} \sim 0.5$, which resulted in the right peak of the histogram being probed and yielded a high number of 118 confirmed UXO out of 131 probed items.

The newly acquired data was then used to retrain the GMM classifier (Step 8) using the features from a precise three-object EMI inversion set. (Note that, since three-target inversions provide seven decay curves per anomaly, only the features closest to the centroids of clusters already identified as containing UXO were considered for GMM training.) The results of the updated GMM-based clustering appear in Fig. 9. A broadening of the histogram peak corresponding to UXO is observed.

Based on the updated histogram, we requested the ground truth from 20 additional suspicious items in order to assess statistically the performance of the classifier. The region containing these items had scores between $-6$ and $-5$. Here, the expert intervened again by visually observing the isosurfaces (and how they encompass the existing UXO clusters) and by considering the spread of the histogram peak corresponding to the UXO. (These tasks can potentially be automated to increase process efficiency). At this stage, if all of the 20 items were returned as clutter, the process would stop, and the scored items would be used to create the final dig list. However, it turned out that 4 out of 20 items were UXO; therefore, the classifier had to be updated once again to ensure that all possible outliers were accounted for.

In order to resolve possible biases from simultaneous treatment of different types of targets, all confirmed UXO items were separated into three categories based on their type.
Fig. 10. The one-component GMM classifier was then applied to features resulting from three-object inversion. (Left) 37-mm targets; Middle: M48 fuzes; (Right) 105-mm targets. The corresponding score histograms are on the bottom row. Initially, 174 anomalies were requested (with 118 being already known) based on the log (score) cutoff value of about \(-6\) for 37-mm targets. After that, 53 anomalies (of which 27 were known) were requested based on a log (score) cutoff value of about \(-5\) for M48 fuzes. Finally, 36 anomalies (18 of them known already) were requested based on a cutoff of about \(-20\) for 105-mm targets.

Fig. 11. Final GMM classifier based on three-object EMI inversion of Camp Butner data. (Left) Three-component GMM classifier score isosurfaces in the classification case based on all UXO targets. (Right) Score histogram showing the number of anomalies scored within a particular range of the log (probability density) in arbitrary units. A total of 377 anomalies were scored as UXO based on a log (score) cutoff of about \(-10\). This value was specified by hand to allow enough statistics and sufficient isosurface separation from identified UXO clusters.

(105-mm, M48, and 37-mm), and each group was used to train a separate one-component GMM classifier that was then used to score every target with a separate score for each target type (Step 10). The ground truth was then selected individually for each of the object types, based on threshold score values that once more were identified visually by the expert. Fig. 10 presents the results obtained with these individual classifiers for 37-mm, M48 and 105-mm target clusters, respectively.

The ground truth obtained as a result of Steps 5–11 comprised 322 requested anomalies, of which 160 were confirmed as UXO. At the final stage, a three-component GMM classifier was trained on the confirmed UXO from the accumulated ground truth and used to score all of the unknown targets (see Fig. 11). A specific threshold was then selected manually, and the final dig list was produced. The results of the overall process are summarized in the ROC curve of Fig. 12. For more details the readers are referred to [60].

IV. CONCLUSION

A hierarchical agglomerative clustering approach combined with Gaussian-mixture probabilistic modeling was applied to a blind UXO test carried out at the former Camp Butner. The ground truth for a total of 322 items was requested in a five-level iterative prediction–correction process. At the end, out of a total of 2291 anomalies, the method yielded 100% accuracy.
in detecting 171 UXO at a cost of 295 false alarms. Machine-learning techniques therefore hold promise for performing high-quality automated UXO discrimination while reducing the workload for human experts and in general speeding up the process.

That said, it would be desirable to improve the process further, so that it results in even better classification and involves the expert even less. One possibility to reduce the number of false alarms and improve the overall quality of the classification may be to incorporate the symmetry exhibited by UXO by analyzing the decay curves of the transverse components of the effective polarizability. It could be also possible to create a library of features from previous discrimination tests. This knowledge, which is already available, could be used to make educated guesses that would help sort out and identify TOI in new live sites. This functionality would be also of help in identifying targets of unknown or unexpected type, which are bound to appear in real sites and confuse any statistical classification algorithm. Even in those cases, we would have two crucial pieces of information: The fact that all UXO exhibit cylindrical symmetry to some extent and the fact that such anomalous targets will be at most a small fraction of the total. The first fact should allow the system to point out good candidates for the expert user to inspect visually, and the second guarantees that this part of the process is not too onerous.

Further improvement of our combined clustering/GMM algorithm must involve ways to perform optimal data clustering, scoring, and thresholding automatically. While external inputs from an expert are valuable for guiding the learning process, it is desirable to reduce human judgment to the point where it provides only systemwide quality checks and classification control. An ideal learning mechanism would first exhaust the information contained in the data themselves before leaving the crucial decision-making to human experts, who continue to be better than computers at some tasks involving pattern recognition, matching, or classification. Such a combined framework may result in overall improved performance and effective resource allocation. A system needing less human input will be also more useful in situations where the operators lack experience, and the system can potentially help them gain it.

REFERENCES


Alex Bijamov received the B.S. degree in physics from Tbilisi State University, Tbilisi, Georgia, in 2006 and the Ph.D. degree in engineering sciences from Dartmouth College, Hanover, NH, USA, in 2011.

From 2011 and 2012, he was a Research Associate with Thayer School of Engineering, Dartmouth College, where he developed advanced models for electromagnetic-induction-based target detection and classification, and worked on statistical signal processing. His research interests include remote sensing, imaging technologies, computer vision, and machine learning.

Dr. Bijamov has served as a Reviewer for the Optical Society of America.

He received a Bronze Medal at the Third Asian Physics Olympiad and a Gold Medal at a Gold Medal at the 33rd International Physics Olympiad.

Juan Pablo Fernández received the Physics degree from Universidad de los Andes, Bogotá, Colombia, in 1990; the Ph.D. degree in theoretical low-temperature physics from the University of Massachusetts, Amherst, MA, USA, in 2004; and the M.S. degree in engineering sciences from Dartmouth College, Hanover, NH, USA, in 2007.

He has taught physics, mathematics, astronomy, Spanish, and scientific writing at several universities in Colombia and the U.S. and has worked as a writer, an editor, and a translator. From June 2010 and May 2012, he was a Visiting Research Scientist with the Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development, Berlin, Germany. He was a Research Associate with the Thayer School of Engineering, Dartmouth College, and a Research Consultant with the U.S. Army Cold Regions Research and Engineering Laboratory, Hanover, NH. He is currently a Signal Processing Content Specialist with MathWorks Inc., Natick, MA, USA.

Dr. Fernández is a member of the Applied Computational Electromagnetics Society. He received a “Francisco de Paula Santander” grant for creative writing from the Instituto Colombiano de Cultura in 1990 and was shortlisted for the Premio Nacional de Literatura in 1998. He also received the Eugene M. Isenberg Award from the Isenberg School of Management, University of Massachusetts, in 1997 and 1998.
Benjamin E. Barrowes (M’99) received the B.S. and M.S. degrees in electrical engineering from Brigham Young University, Provo, UT, USA, in 1999, and the Ph.D. degree from Massachusetts Institute of Technology, Cambridge, MA, USA, in 2004.

From 2004 to 2005, he was a Director’s funded Postdoctoral Researcher with the Physics Division, Los Alamos National Laboratory, Los Alamos, NM, USA. He is currently a Physicist with the U.S. Army Cold Regions Research and Engineering Laboratory, Hanover, NH, USA. His research interests include electromagnetic wave theory and modeling with applications on wind–wave interaction, electromagnetic scattering from the sea surface and from random media, nanoscale energy-generating techniques, computer interface methodologies, electromagnetic induction models for nonspherical geometries, and biological electromagnetic phenomena. His other research interests include automatic code conversion/translation and arbitrary precision computing.

Dr. Barrowes received two Rocky Mountain Space Grant Consortium GRANTS, and a National Science Foundation Graduate Fellowship Award.

Irma Shamatava received the M.S. degree in radio physics from Sokhumi State University (formerly the Sukhumi branch of Tbilisi State University (TSU)), Tbilisi, Georgia, in 1994 and the Cand.Sci. degree in radio physics (applied electromagnetics) from TSU, Tbilisi, in 1997.

In 1994, he was a member of the Research Staff of the Laboratory of Applied Electrodynamics, Department of Physics, TSU. At the same time, he was with the Department of Physics and Mathematics, Sokhumi State University, as a Senior Teacher and became an Associate Professor in 1998. From 1998 to 1999, he was a Postdoctoral Fellow with the National Technical University of Athens, Athens, Greece, performing research on computer simulation of electrostatic discharge; electrodynamic aspects of electromagnetic compatibility, numerical modeling of conformal antennas; electromagnetic wave scattering; field visualization and identification of objects by scattered field analysis; investigation of wave propagation through anisotropy, plasma, and chiral media; and innovative numerical methods. From June to August 2005, he was a Visiting Scientist with the Department of Earth and Ocean Science, University of British Columbia, Vancouver, BC, Canada. From April 2007 to May 2013, he worked with Sky Research, Inc., Hanover, NH, USA. He is currently a Research Associate Professor with the Thayer School of Engineering, Dartmouth College, Hanover, NH, USA. His current research interests include numerical modeling of forward and inverse electromagnetic scattering by subsurface metallic objects, investigation of global magnetoequasi-static fields due to ionospheric currents, and electrostatic fields’ effects on robotic and construction systems for Mars and Lunar exploration, magnetic nanoparticle hyperthermia for cancer treatment, and advanced sensors for DNA sequencing.

Frilon Shubitidze (A’98–M’04–SM’06) received the M.S. degree in radio physics from Sokhumi State University [formerly the Sukhumi branch of Tbilisi State University (TSU)], Tbilisi, Georgia, in 1994 and the Cand.Sci. degree in radio physics (applied electromagnetics) from TSU, Tbilisi, in 1997.

In 1994, he was a member of the Research Staff of the Laboratory of Applied Electrodynamics, Department of Physics, TSU. At the same time, he was with the Department of Physics and Mathematics, Sokhumi State University, as a Senior Teacher and became an Associate Professor in 1998. From 1998 to 1999, he was a Postdoctoral Fellow with the National Technical University of Athens, Athens, Greece, performing research on computer simulation of electrostatic discharge; electrodynamic aspects of electromagnetic compatibility, numerical modeling of conformal antennas; electromagnetic wave scattering; field visualization and identification of objects by scattered field analysis; investigation of wave propagation through anisotropy, plasma, and chiral media; and innovative numerical methods. From June to August 2005, he was a Visiting Scientist with the Department of Earth and Ocean Science, University of British Columbia, Vancouver, BC, Canada. From April 2007 to May 2013, he worked with Sky Research, Inc., Hanover, NH, USA. He is currently a Research Associate Professor with the Thayer School of Engineering, Dartmouth College, Hanover, NH, USA. His current research interests include numerical modeling of forward and inverse electromagnetic scattering by subsurface metallic objects, investigation of global magnetoequasi-static fields due to ionospheric currents, and electrostatic fields’ effects on robotic and construction systems for Mars and Lunar exploration, magnetic nanoparticle hyperthermia for cancer treatment, and advanced sensors for DNA sequencing.

Dr. Shubitidze and his group received the Munitions Response Project-of-the-Year Award given by the Strategic Environmental Research and Development Program in 2011.