INDEPENDENT ANALYSIS OF DETECTION DATA FOR UNEXPLODED ORDNANCE FROM THE REMOTE IDENTIFICATION SYSTEM (REMIDS)

HOLLIS H. (JAY) BENNETT, JR.
U.S. Army Engineer Waterways Experiment Station
3909 Halls Ferry Road
Vicksburg, MS, USA 39180-6199
Phone Number: (601) 634-3924
Facsimile Number: (601) 634-2732
E-mail Address: benneth@mail.wes.army.mil

THOMAS W. ALTSHULER, LISA PORTER,
FRANK ROTONDO, DAVE SPARROW
Institute for Defense Analyses
1801 North Beauregard Street
Alexandria, VA, USA 22311
Phone Number: (703) 578-2715
Facsimile Number: (703) 578-2877
E-mail: taltshul@ida.org

ABSTRACT

The information presented describes the analysis of data collected from a helicopter-mounted multispectral line scanner system with respect to the detection of unexploded ordnance (UXO) at the terrain surface. The system, REMote IDentification System (REMIDS), consists of an active/passive multispectral line scanner, post-flight processing and display equipment, and navigational equipment. The use of REMIDS minimizes the risk to personnel during the environmental assessment and analysis of the site.

The data used in the analyses are from Yuma Proving Ground (YPG), Arizona and Fort Rucker, Alabama. The items used at the sites consist of inert items of the following types: mines, grenades, rockets, mortars, projectiles, and bombs. The performance of the system with respect to the detection of the different classes of items will be presented. Also, the performance of the system based on background clutter will be evaluated. The adaptation of this remote identification system will aid in the cleanup of Department of Defense (DoD) sites with UXO contamination.

The REMIDS collects three channels of optically aligned image data consisting of two active laser channels, one polarized reflectance and the other total reflectance, and one passive thermal infrared channel. The system also incorporates onboard sensor data recording and post-processing insertion of differential Global Positioning System (GPS) coordinates. The system can be flown at various altitudes and forward speeds to characterize sites for the presence of surface UXO. The detection is based on the remote identification of surface anomalies and materials that

indicate the presence of surface UXO contamination. The system evaluation is funded by the Environmental Security Technology Certification Program (ESTCP).

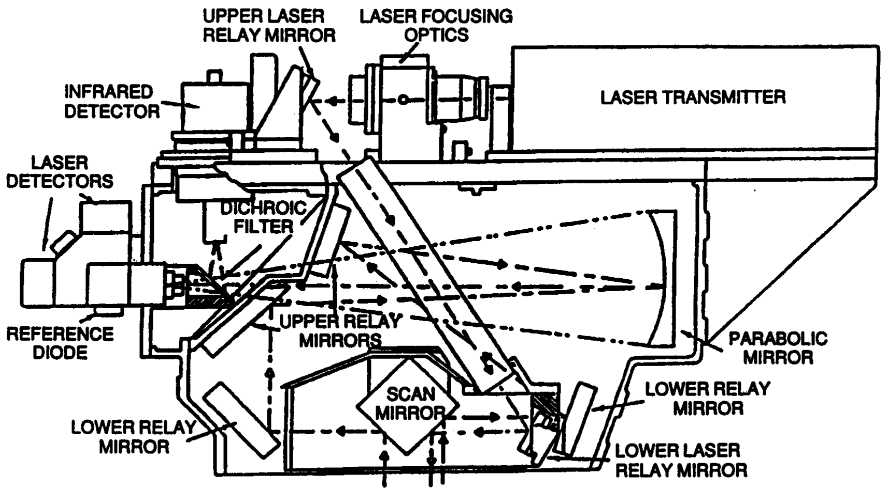
INTRODUCTION

The Institute for Defense Analyses (IDA) was requested by ESTCP to perform an independent evaluation of the data collected by REMIDS. This paper describes the REMIDS developed by the U.S. Army Engineer Waterways Experiment Station to detect surface UXO, and reviews the performance of REMIDS in tests at the Yuma Proving Ground (Arizona) and at Ft. Rucker (Alabama). The principle behind REMIDS is to enhance the discrimination of surface UXO by relying on multiple signatures: surface UXO may exhibit a unique combination of reflectance, polarization, temperature, and footprint (shape), as compared to natural objects in the UXO's surroundings. Discrimination based on four signatures is in principle greater than that based on fewer signatures.

SYSTEM BACKGROUND

The airborne data collection system consists of an active/passive line scanner, real-time processing and display equipment, and navigational equipment and is described in detail elsewhere (Ballard 1992). The scanner collects three channels of optically aligned image data consisting of two active laser channels (one polarized reflectance and the other total reflectance) and one passive thermal infrared channel. The real-time processing and display system is based on a massively parallel processor. The system has a scan rate of 350 scans per second with 710 data pixels per scan. The area that can be covered in 18 flight hours will depend upon the profile of the data collection flights (i.e., altitude and forward speed), as well as the chosen flight path. For example, if the 0.5 km by 1 km site at Yuma is traversed crosswise (i.e., using 500 m passes), then during one flight hour, only about 15 minutes is dedicated to actual data collection. The rest of the time is spent making turns and lining up for the next pass. In this case, with a pixel resolution of 1.9" by 1.9", one flight hour yields about 100 acres of coverage. On the other hand, if a site is traversed lengthwise, the coverage will be much greater; estimates provided by the REMIDS team are of order 300 acres per flight hour. A lower resolution will also allow for more coverage per flight hour; decreasing the resolution to 4.5"x 4.5" will yield about a factor of five increase in area coverage per flight hour. The detection is based on the remote identification of surface anomalies and materials that indicate the presence of surface UXO contamination. A cut-away diagram of the scanner is shown in Figure 1.

Figure 1. Scanner Physical and Optical Layout



PHENOMENOLOGY

The REMIDS surface mine/UXO detection algorithm is comprised of three sequential steps. In the first step, the reflectance, polarization, and thermal returns are used to assign a classification to each pixel. Such a classification provides information on material type, and allows for some reduction in the total number of pixels that must be investigated. In the second step, pixels of like classification are joined together to form objects. Each object is characterized by its total area, boundary, and distance between its two furthest pixels. At the end of the second step, a table is generated that groups the objects together according to location and gives both the spectral and size/shape information of each object. In the third step, the operator decides which objects are targets, using the table provided in the second step to guide his decisions. In the following, we discuss each of these three stages in some detail, and we provide a quantitative estimate of the role of each step in reducing the false alarms.

1. Stage 1: Spectral Discrimination

The first step in the algorithm, which we will call "Stage 1", relies on the fact that different materials will yield different polarization, reflectance, and thermal responses. The reflectance is a function of both the refractive index of the object and its orientation relative to the incident light. Metals, in general, will have a very high reflectance because they have a complex refractive index, which essentially means that they do not permit the penetration of an electromagnetic wave to any significant depth and thus reflect almost all of the incident light. (In the limit of a perfect conductor, the penetration depth goes to zero and the reflectivity goes to unity.)[1]

---

[1] To be more precise, the reflectivity of a metal is very high as long as the frequency of the light is below the plasma frequency of the metal. Aluminum will reflect light at all optical wavelengths.

The polarization of a material is essentially a measure of its "smoothness", which according to the Rayleigh criterion, is proportional to $\lambda / (h_{rms} \cos\theta)$, where λ is the wavelength of the incident light, θ is the angle of incidence, and h_{rms} is the RMS height of the surface. A material can therefore appear "rough" and hence yield a low polarization return at 1.06 μm, and yet appear "smooth" and yield a high polarization return at 10.6 μm. Hence, one cannot predict the polarization return of a material based on how it looks or feels. One example of this is rust, which one might intuitively expect to have low polarization, but which has been found to yield a significant polarization return at 1.06 μm.

The thermal returns of both ordnance and background will depend upon their thermal properties as well as on the time of day and the weather conditions. The discrimination capabilities of the thermal channel should be best soon after sunrise and sunset, when objects with different thermal diffusivities and surface absorptivities and emissivities will heat up and cool down at different rates. Near mid-day, when thermal equilbrium is approached, the thermal discrimination between objects will be more difficult. This will be shown in the performance section for the Ft. Rucker data, where it will be clearly seen that the discrimination capabilities of the thermal channel at 0731 a.m. are far superior to those at 11:22 a.m.

The classification of the test site pixels according to their spectral information is a complex process. For each pixel in question, a set of distances, D_i, are defined for each of the possible target material classifications,

$$D_i = \sqrt{\left(\frac{P - \langle P_i \rangle}{\sigma_{P_i} w_{P_i}}\right)^2 + \left(\frac{R - \langle R_i \rangle}{\sigma_{R_i} w_{R_i}}\right)^2 + \left(\frac{T - \langle T_i \rangle}{\sigma_{T_i} w_{T_i}}\right)^2}\ , \qquad (1)$$

where *i* is the ith classification, angular brackets indicate the mean, and standard deviations are denoted by σ. The means and standard deviations are obtained from the pixels comprising the targets in the calibration site, and the classifications include only potential target materials. The parameters w_{Pi}, w_{Ri}, and w_{Ti} are weights that account for the fact that the pixels in the calibration site are not necessarily representative of those in the test site. For example, one should expect much more variability in the thermal channel for the test site than for the calibration site due to time drift. Furthermore, in order to reduce the complexity, only a few classifications are chosen, which means that several different materials will fall into one general class, implying greater variability in all three channels for the test site than for the calibration site. Typically, at Yuma, the values of w_{Pi} were chosen to be close to unity because the values of s_{Pi} were large; those of w_{Ri} were around two, and those of w_{Ti} were at least two, if not greater.

For each test site pixel, the smallest D_i, D_{i_min}, determines its target material classification, if D_{i_min} is less than or equal to a threshold value, D_{thresh}. If D_{i_min} exceeds D_{thresh}, then the pixel is classified as background. D_{thresh} is an adjustable parameter, but it is assumed to be the same value for each material classification. On the other hand, the quantities w_{Pi}, w_{Ri}, and w_{Ti} may vary from one classification to another. Thus, the total number of adjustable parameters is

3N+1, where N is the number of target material classifications. We note that by fixing the 3N values of w_{Pi}, w_{Ri}, and w_{Ti}, and varying only D_{thresh}, one can build up a ROC curve.

The 3N+1 adjustable parameters are first chosen to optimize the results from the calibration site. This in itself is a non-trivial task and can take as much as one work day to complete. If it is then found, when equation (1) is evaluated for the pixels of the test site, that something does not seem right (e.g., that pixels covering a very large area are all classified as being the same target material), the operator will vary some of the 3N+1 values and repeat the process. Thus, the calibration site provides the operator with initial values for each of the adjustable parameters, but several iterations through the test site data are usually necessary to obtain the optimal set of parameters.

In Table 1., we provide the means and standard deviations of P, R, and T obtained from a noontime flight at Yuma Proving Ground on June 26, 1996. The first column is the number of pixels used to generate the statistics. Those items marked with asterisks comprise the target material classifications used in the evaluation of the test site data. Their means and standard deviations were obtained from the calibration site, except for the dielectric mines, whose means and standard deviations were taken from mines in the test site, as none were present in the calibration site.[2]

Two comments need to be made regarding Table 1. First, it may seem surprising that the reflectivity of aluminum is lower than many of the other entries. This is because the aluminum targets in the calibration site that were used to obtain these numbers were very small. Thus, edge pixels, which were affected by the background, had a significant effect on the mean. Large sheets of aluminum, conversely, yield strong polarization and reflectance returns. Second, two entries warrant further clarification. Desert pavement refers to a rocky blackened crust that covers much of the test site at Yuma and which appears quite smooth at 1.06 μm, yielding fairly high polarization returns. Another background feature yielding even higher polarization returns than the desert pavement were rocks covered with a glass-like sheen, called desert varnish. These rocks were located mainly in the wash areas of the test site.[3] At some sites, such as Ft Rucker, the polarization channel alone can provide very good distinction between the targets and the background, but at Yuma, this is clearly not the case. The additional information provided by the thermal channel does allow for the desert pavement to be fairly easily separated from potential targets, but the desert varnish is very difficult to distinguish from the iron/olive drab paint class, even with the information from all three channels.

---

[2] One might wonder how they were able to classify objects as dielectric mines, given that there were none in the calibration site. It turns out that it was known that one dielectric mine was located at the corner of the test site. This one mine provided about 25-30 pixels of information on P, R, and T. Although the standard deviation on 25-30 pixels is large, this provided a starting point for the analysis of subsequent objects. A few additional objects with similar returns were then picked up and added to the database for dielectric mines.

[3] A wash area is an area where flash floods caused a lot of water of flow through the area. These areas represented roughly 30% of the site.

Table 1. Polarization, Reflectance, and Thermal Returns at Yuma

Material	# Pixels	<P>	s_P	<R>	s_R	<T>	s_T
desert pavement	2.1e+5	149.91	9.83	82.27	9.24	148.82	15.14
sand and fines	2523	105.17	7.18	166.00	7.97	129.11	5.45
desert varnish	189	180.57	15.27	92.58	12.19	108.93	13.66
iron (oxidized) and olive drab paint *	192	206.98	35.22	58.72	22.43	98.96	17.62
white paint*	15	197.07	36.69	145.67	22.60	73.20	17.50
aluminum*	80	223.30	31.98	101.85	27.95	76.48	33.53
dielectric mines (white plastic)*	180	49.58	20.72	131.66	13.60	90.67	15.93

We note that completion of Stage 1 of the REMIDS algorithm does not require that the helicopter obtain the calibration data, land, and then obtain the target data. All polarization, reflectance, and thermal data are post-processed. Thus, the REMIDS approach is to fly over the calibration site and then fly directly to the real site. In fact, this is necessary for thermal channel calibration data to have any applicability.

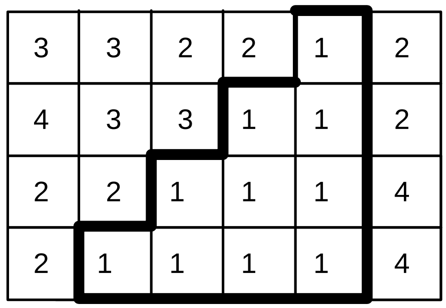
The completion of Stage 1 results in a significant reduction of the total number of pixels that might contain targets of interest. Nonetheless, Stage 1 cannot always be relied upon to provide sufficient discrimination from the background to be a viable surface/UXO mine detection system in and of itself. This is particularly true at sites such as Yuma, where the existence of desert varnish resulted in many tagged pixels. Other sites may fare better; for example, as can be seen in Table 1., sand is easily distinguishable from the targets. Grass, too, is an "easy" background, as will be seen in the results of the Ft. Rucker test, presented in the performance assessment section of this report.

In addition to being site-dependent, the reduction in total pixels of interest after Stage 1 is also dependent on the choice of D_{thresh}: a larger D_{thresh} means that more pixels will be tagged as possible targets. The optimal value of D_{thresh} is typically determined after a few iterations through the entire algorithm. Usually, the operator will begin with a conservative value of D_{thresh} and focus in on a few items that he believes are targets, and that are somewhat difficult to pick out. He will then decrease D_{thresh} until those objects are no longer cued. At Yuma, a conservative value of $D_{thresh} = 1.6$ resulted in a reduction in the number of tagged pixels after Stage 1 by about two and a half orders of magnitude, from about three hundred million pixels (the total overflown area) to about six hundred thousand.

2. Stage 2: Size/Shape Discrimination

In the second step of the REMIDS algorithm, which we will call "Stage 2", pixels of like material classification are joined together to form objects. Figure 2. illustrates an example in which 10 pixels of material class 1 are combined to form an object with an area of 36.1 inches, a boundary of 30.4 inches, and a length, defined as the distance between the two furthest pixels, of 10.75 inches.

Figure 2. Joining objects of like classification



Software has been developed which performs this step automatically. The operator has the option of specifying how he wants the objects to be formed; for example, he may specify that pixels must be contiguous, or that they need not be. Furthermore, the operator can customize the software so that, in addition to forming objects from tagged pixels, it also filters out many of the objects based on their size and/or shape. An object with a very large area, for example, may be rejected, as may one that is too small, or too thin, or very irregularly shaped.

The level of discrimination provided by the size/shape filter will greatly affect the number of objects that are passed on to the operator: too little pre-screening in Stage 2 will result in too many man-hours spent in the final step, but too much pre-screening will result in too many missed targets. Just as with the 3N+1 spectral parameters in Stage 1, the optimal level of discrimination used in Stage 2 is ultimately determined after several iterations through the data. At Yuma, for example, it was decided that a minimum size be set such that any object smaller than four pixels was rejected. This served to greatly reduce the number of false alarms due to desert varnish, at the expense of missing targets smaller than four pixels, such as grenades and Valmeira mines. On the other hand, discrimination based on "irregular shape" was found to be unreliable at Yuma because variabilities in helicopter air speed caused oblong targets to appear irregular in shape. Thus, the only discrimination provided by the size/shape filter in Stage 2 for the Yuma data was that of the minimum size requirement.

The output of the size/shape filter in Stage 2 is a table in which objects are ordered by location (image and scan line). For each object, the material classification and shape information (area, boundary, length) are given. At Yuma, for the iteration in which D_{thresh} was set to 1.6, where 600,000 pixles were passed to the size/shape filter, approximately 20,250 candidate objects were selected. Of these, roughly 20,000 were not targets.

3. Stage 3: Operator Discrimination

The table provided to the operator by the shape/size filter allows him to decide which images upon which to focus his attention. The operator chooses a spectral channel (usually the polarization channel) and scrolls through the images, viewing several suspicious objects at once. Usually, the man-made ordnance items are easily discernable by eye.[4] Thus, a trained operator is able to identify many of the ordnance items and eliminate many of the false alarms very quickly, without having to study each individual object in great detail. For example, at Yuma, about 95% of the roughly 20,000 false alarms were located in the wash areas, which is where most of the desert varnish was located. After carefully examining a few of these objects, the operator realized that most of the items located in these areas were false alarms. Thus, the operator was able to scroll through about 19,000 objects fairly quickly, without having to focus on any of them in detail.[5] Those hundred or so objects that the operator cannot identify or eliminate right away are carefully examined in each of the three spectral channels. We note that the operator not only determines whether a suspicious object is a target, he also classifies the target if he decides that it is an ordnance item.

It is clear that the role of the operator is crucial to the success of this system. The operator is involved in each of the three stages, and the total time involved is dictated by his efficacy and skill. For the data collected at Yuma, one day was devoted to processing the calibration site data and choosing the initial values of the 3N+1 spectral parameters. The analysis of the test site data then took three days, the vast majority of which was due to time spent by the operator. Of those three days, roughly 60% of that time was spent scrolling through images, while the remainder of the time was dedicated to the detailed examination of the roughly one hundred suspicious objects that could not be quickly identified.

Given that an area of only 0.5 km^2 entailed three man-days of tedious data analysis, an obvious question to address is whether the role of the operator can be automated. Clearly, it should be possible to reduce the amount of time spent scrolling through images, since that is determined primarily by the level of pre-screening provided by the size/shape filter. At Yuma, this pre-screening was limited to just a minimum size requirement. It is believed by the developers of this system that a robust size/shape filter can be developed that will allow for extensive filtering of the objects before they are passed to the operator, but that has yet to be tested. However, if one opts to rely heavily on a size/shape filter to screen the objects, then one should expect some degradation in performance over that of the current system, because it is unlikely that any computer can provide the level of discrimination of the human eye. As a specific example, it is likely that many more false alarms would be declared in the wash areas at Yuma, because the

---

4 While man-made ordnance is fairly easily distinguished from natural backgrounds, it should be noted that it may not be easily separated from man-made clutter. The test site at Yuma was a fairly "clean" site, with very little anthropic clutter.

5 There were some ordnance items located in the wash area as well. Some of these targets were discernable based on their size, while others could be distinguished by their spectral signature. Furthermore, the rocks often possessed irregular shapes compared with the man-made targets. However, some targets were lost; small items made of iron or covered in olive drab paint would have been particularly difficult to pick out.

operator relied in part on the irregular shapes of the desert varnish-covered rocks. These irregularities, although perceptible to the human eye, would not be so easily distinguished by a computer-based algorithm. In addition, as the operator scans the images, he develops an "intuition" for where the targets and false alarms are most likely to be located. That is, humans possess pattern recognition capabilities that may be difficult to replicate in a computer program. Nonetheless, it seems impractical to rely on the operator to perform the bulk of the discrimination for large areas.

LIMITATIONS OF THIS SYSTEM

A key limitation of the REMIDS system is that its performance is highly site-dependent. The grassy background at Ft. Rucker enabled the system to easily pick out the targets based on their spectral information alone; no shape filter was needed. On the other hand, the desert varnish at Yuma rendered the spectral information insufficient. It became apparent from the Yuma data that both the shape filter and especially the target recognition capabilities of the operator will play a critical role in this technology for any backgrounds with spectral returns close to those of the targets of interest.

In addition to the fact that some backgrounds render discrimination based solely upon spectral information impossible, there are other physical limitations of the system. In order for this system to work, the laser must have a line of sight to the surface ordnance items. Thus, broad-leaf vegetation, trees, and understory can all pose a laser penetration problem. Dust-covered ordnance may also be difficult to see; even if the dust-covering is such that some of the surface is still exposed, the classification of the UXO may be difficult. Finally, the system cannot be used when the targets are covered with snow.

The REMIDS system also works much better if one has samples of the types of materials one expects to find at the site so that the spectral classes can be calibrated. The polarization returns for painted objects can vary greatly depending upon the type of paint, for example. Also, the polarization and reflectance returns will vary with ordnance age. For some material types, this is not a big problem; rust, for example, still yields a fairly high polarization return. However, the polarization and reflectance returns of oxidized brass are much lower than those of new brass, and this caused some problems at Yuma. Specifically, snake-eye fin assemblies that were at the Yuma site were made of painted brass. The non-weathered fin was an easy item to detect, but the weathered fin was not associated with any of the material classes and hence went undetected.

If this system were used at a site at which only limited information was available about the types of ordnance items expected, then the performance would likely be below that exhibited at Ft. Rucker and at Yuma. In principle, one could try to include as many material classes as possible; there is no limit to the number of material classes used in the spectral filter. In practice, however, the more classifications, the more complex the processing, since the total number of free parameters used in the spectral filter is $3N + 1$, where N is the number of spectral classifications. Hence, to optimize the performance of this system, one should find out as much as possible about the UXO at the site in question, particularly the material composition and age of the ordnance.

A major lesson learned during the testing of the REMIDS system is the difficulty of 100% site coverage. At Yuma, the helicopter flew over only about 96% of the site, and therefore, a P_d of 100% was not achievable. There are four main factors that determine the amount of site coverage: (1) the amount of overlap between adjacent passes, (2) the direction of the wind gusts, (3) the intensity of the wind gusts, and (4) pilot reaction time to the wind gusts. An increase in overlap does not guarantee full coverage because the wind gusts are very important. It is recommended that a GPS-aided flight guidance system be used, and that those areas that are missed be covered by a second flight with minimal duplication of the coverage.

PERFORMANCE ASSESSMENT

In this section, we summarize the performance of the REMIDS system at two test sites: Ft. Rucker, AL and Yuma, AZ. It will be seen that the performance at Ft. Rucker was superior due to the fact that the background was mainly grass. In essence, Ft. Rucker looked like a "short rough" on which the targets were easily seen, even with just the polarization information. At Yuma, the existence of desert varnish, with its high polarization return, made the targets much more difficult to distinguish, and a 4-pixel minimum object size was required in the analysis. This resulted in a limitation of the size of the objects that could be detected at Yuma (e.g., grenades fall below this size limit).

A. FT. RUCKER RESULTS

We present the performance curves in Figures 3.1. Fig. 3.1(a) shows the performance of the individual spectral channels with respect to the detection of aluminum items, while Fig. 3.1(b) shows the performance of the individual spectral channels with respect to ferrous and painted surface items. The performance of the combined spectral channels for the detection of both aluminum and ferrous/painted targets is shown in Fig. 3.1(c). Fig. 3.1(a) shows that the polarization channel alone was sufficient to detect the aluminum ordnance items: 100% P_d was achieved with only 15 false alarms in 2400 m^2. In Fig. 3.1(b), it is seen that the performance of the polarization channel in detecting ferrous or painted objects, while not as remarkable as for aluminum, was still quite good, with a P_d of about 95% at 15 false alarms. It is clear from Fig. 3.1(c) that the combination of the three spectral channels yielded excellent performance: greater than 80% P_d was achieved with essentially zero false alarms, while 100% P_d was possible with only 20 false alarms. Thus, the spectral filter alone was sufficient at Ft. Rucker; no size/shape filter was required, and the operator's role was minimized.

Fig. 3.1(a): ROC curve for Ft. Rucker: Aluminum Targets Only

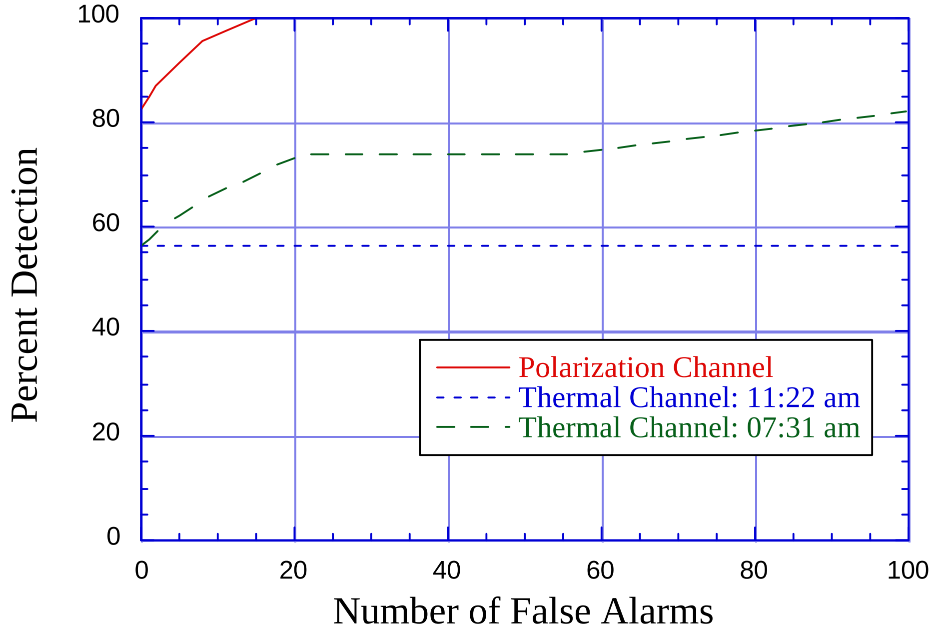


Fig 3.1(b): ROC curve for Ft. Rucker: Ferrous and Painted Targets Only

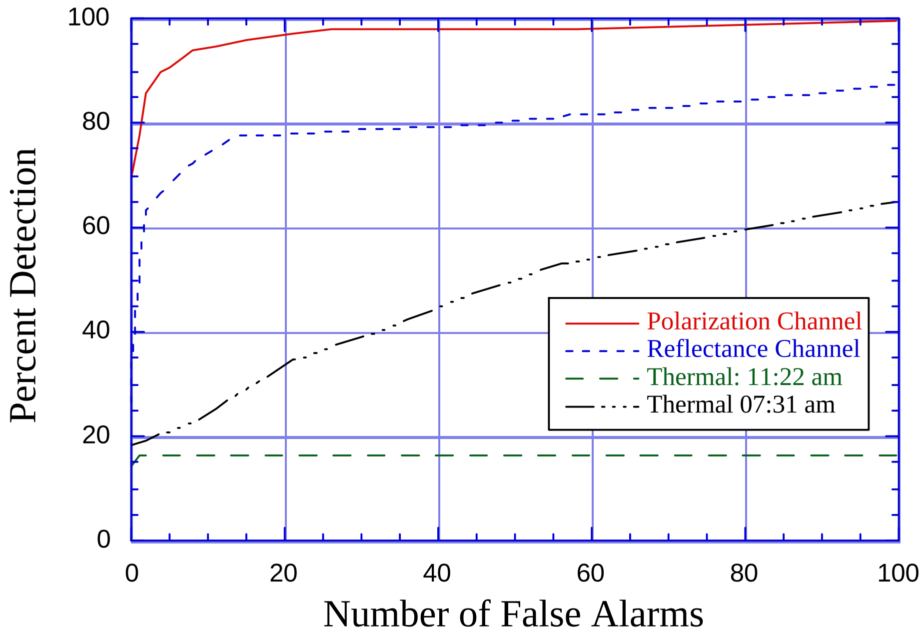
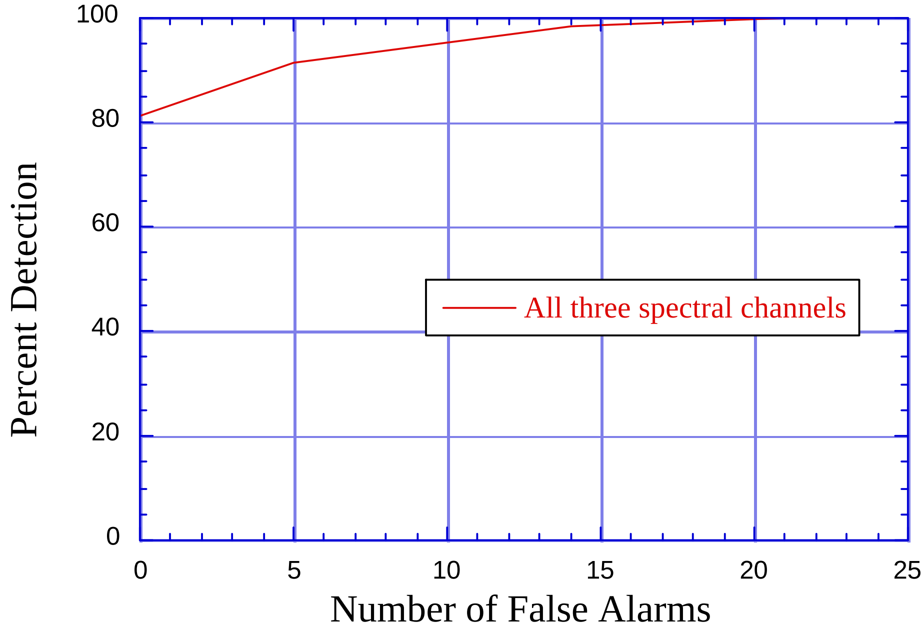


Fig. 3.1(c): ROC curve for Ft. Rucker: All Targets



B. YUMA RESULTS

The performance at Yuma is somewhat difficult to assess quantitatively for two reasons. First, insufficient information was provided about target locations. Documentation consisted solely of the contents of each 25 m x 25 m cell, with a rough estimate of the target location(s) within each cell, such as "center" or "northeast corner." For some cells, a hand-drawn sketch was provided. Clearly, the location accuracy of the REMIDS system could not be tested, and detections could not be correlated to targets in cells containing multiple items.

The second, and much more unsettling, reason for the difficulties in evaluating the REMIDS performance at Yuma was that there were uncertainties in the baseline. After the test was conducted, the REMIDS team walked much of the site, taking notes on target types and locations within each cell. They found that, for some of the cells, there seemed to be significant discrepancies between what was documented as the baseline, and what they actually observed. It was learned that the site is somewhat "dynamic"; that is, items are "borrowed" and moved around. The documentation provided at that time was based on a test site survey done in November of 1994. The site manager agreed to provide an update of the site baseline, which was provided in March 1997. The problem is that it is not known for certain how much the site changed between June 1996, when the test data was taken, and March 1997, when the update to the baseline was determined. By comparing the baseline information provided by the Yuma site manager for November 1994 and for March 1997, we found that, of the 132 cells that actually contain some type of ordnance, 20 cells show significant differences between the two time periods. It is not known for certain whether the contents of all 20 cells had actually changed, or whether the documentation for some of those cells was incorrect. Either way, however, those 20 cells must be considered questionable.

The REMIDS team attempted to grade themselves as best they could, given these two limitations. However, because they had walked much of the site, and had been provided with the November 1994 and March 1997 baselines, it was clearly impossible for them to grade themselves "blindly". Thus, we at IDA have attempted to grade the REMIDS system as objectively as possible. We describe our methodology below.

First, in consult with our sponsor, we decided to limit our focus to those targets (1) that are larger than the four-pixel threshold, and (2) that are of interest to the UXO community. The targets chosen were: 500 lb bombs, 2.75" rockets, 81 mm mortars, 105-mm projectiles, and 155-mm projectiles. Dielectric mines, although not of interest to the UXO application that was being examined, were also included because they were a particularly easy target for REMIDS to detect,[6] and thus provided a means of estimating the location accuracy of the system. Excluded items included grenades, Valmeira mines, gator mines, volcano mines, and painted mines. All but the last of these were too small to be detected with the four-pixel threshold.

For each of the categories of interest, we determined for which of the cells the documentation in November 1994 and March 1997 was consistent and clear. In October 1997, we submitted these cells (59 cells altogether, containing 186 targets) to the site manager and requested that he obtain GPS data for each item in each cell.

However, we discovered upon receiving the requested GPS data in November of 1997 that the number of items in 6 of the 59 cells differed from what was documented in March 1997. We had chosen those 59 cells specifically because they appeared to remain constant over time, according to the 1994 and 1997 documentation. Thus, the fact that the contents of some of the cells seemed to have changed in less than a year caused some concern as to whether the GPS coordinates provided for the remaining 53 cells were relevant to the June 1996 test. (Certainly, if the items in some cells had changed, the items in others may have been moved around.)

In February 1998 the site manager attempted to determine what had happened in those six cells. He concluded that, while two of the six cells seemed to have acquired additional items, for the four cells which had fewer items surveyed than what was documented to be there, the original documentation was correct, and the surveyors simply had missed some items. We decided to discard only the two cells where the number of items had seemingly increased. By keeping the four other cells, we may slightly overestimate the probability of detection, but we estimate that the effect of this error will be less than 1%.[7] If we assume that the GPS coordinates of the items in the remaining 57 cells have not changed since June 1996, then the data set is reduced to a total of

---

[6] Recall Table 1. and the fact that the dielectric mines gave a polarization return that easily distinguished them from all other targets and backgrounds.

[7] For example, say a cell had 6 items in it, but only 5 GPS coordinates were provided. By grading this cell, we are theoretically giving the REMIDS team 6 chances at 5 objects. However, the real effect will be minimal if we assume that the items are randomly placed in the 25m x 25 m cell.

168 targets. However, one additional cell was not used because inadequate information was provided by the REMIDS team.[8] Thus, the useful data set was reduced to 167 targets in 56 cells.

To further complicate matters, the REMIDS system experienced problems with its GPS system for 13 of the 59 cells. Thus, when one excludes the cells in which no GPS data was collected by REMIDS *and* the cells discussed above, one obtains a total of 141 targets in the baseline (43 cells).[9]

We used the GPS data for the dielectric mines and the 500 lb bombs to estimate the location accuracy of the REMIDS system, because these are considered easy targets for the REMIDS system. There were a total of 17 mines and 2 bombs that were part of the final target baseline. However, the REMIDS system missed 2 of the 17 mines. On the other hand, the REMIDS system often overflew the same target twice, and therefore, for several of the 17 detected mines and bombs, there were two sets of GPS coordinates provided by REMIDS, resulting in a total of 27 data points that were used in the calculation of the location accuracy. We determined the location accuracy of the REMIDS system to be 1.55 meters in the easting direction and 2.18 meters in the northing direction. (The precision of the GPS system used at the site was 0.1 m, according to the site manager.) These values lend support to the assumption that the GPS coordinates in the cells whose contents presumably have not changed are relevant to our analysis. In addition, we found that there was an offset bias in the REMIDS position compared to ground truth, specifically a 1.12m offset in the easting direction and 0.44m offset in the northing direction. We left these offsets in the data during our analysis of the probability of detection outlined below, because this was the data as derived from the test. However, we will also show that correcting for these offsets only marginally changes the device's detection efficiency.

We then estimated the probability of detection of the REMIDS system, as a function of radius surrounding the target. To perform this calculation we excluded the dielectric mines, which are not relevant to UXO clearance. Thus, the set of ground truth target data against which REMIDS was graded consisted of 124 ordnance items (70 81mm mortars, 24 105 mm shells, 15 155mm shells, 13 2.75" rockets, and two 500 lb bombs), distributed amongst 26 cells. In the remaining discussion, "targets" refer to ground truth ordnance items emplaced at Yuma, while "candidates" refer to REMIDS detections. The analysis flow was structured as follows:

1) Any REMIDS candidate that was both the closest candidate to a target and also within a critical radius, R_{crit}, of the target was matched as a possible detection of the target. This was computed for all targets.

2) A candidates passing the requirement (1) was considered a detection of a target if it was closer to that target than to all other targets. The matched candidate and target were then removed from their respective lists.

---

[8] The cell contained one dielectric mine and two painted mines on top of a 4' x 4' aluminum plate. The REMIDS team noted only that there were three objects on top of a plate, and gave only the GPS coordinates of the center of the plate. Because the dielectric mines were used only to determine the location accuracy of the REMIDS system, it was decided that this cell was not needed.

[9] One of those 13 cells is also one of the 4 cells where the March 1997 documentation indicated more items than what was surveyed in November 1997. Hence only 3 cells having fewer surveyed items than what was documented were in the final target baseline.

3) Steps (1) and (2) were repeated for the reduced lists of candidates and targets until there were no more targets or there were no more candidates within R_{crit} of the remaining targets.

4) The detection probability of REMIDS as a function of R_{crit} was computed as the number of detected targets for a given R_{crit} divided by the total number of targets (which is 124 in this analysis).

5) In addition to the detection probability, ordnance identification could also be tested in this analysis (the REMIDS team used imagery from the polarization channel). The fraction of the detected ordnance that was correctly identified was measured as a function of R_{crit}.

The results of this analysis are summarized in Table 2. and Figure 3.2. Table 2. gives the breakdown of detection probability versus R_{crit} for each ordnance type. These results are summarized in Figure 3.2 which shows the probability of detection as a function of R_{crit} for the combined set of ordnance items. Using Figure 3.2 and taking an example value of R_{crit} = 5m (which is more than a 2 sigma cut according to our estimate of the system's position resolution), we find that REMIDS detected 76% of the target test sample and correctly identified 95% of the ordnance detected.

Table 2. Probability of Detection by Ordnance Type

ordnance type	Pd@Rcrit=1m	Pd@Rcrit=3m	Pd@Rcrit=5m	Pd@Rcrit=7m
81mm	14.3%	57.1%	72.9%	78.6%
105mm	4.1%	45.8%	75.0%	75.0%
155mm	0.0%	46.7%	86.7%	86.7%
2.75"	0.0%	30.8%	76.9%	92.3%
500lb	0.0%	50.0%	100.0%	100.0%

As discussed above, there was an offset bias in the REMIDS position measurement. In order to determine how an offset correction would affect the detection probability, we ran a Monte Carlo that simulated the device's easting and northing position resolutions, both with the offsets and with a correction that removed the offsets. We found that correcting for the offset would improve the detection efficiency by only 2.5% over the uncorrected sample for the case when R_{crit}=5m. Thus, REMIDS would have achieved about a 78% detection probability at R_{crit}=5m had the operators corrected the offset bias at the calibration site. [10]

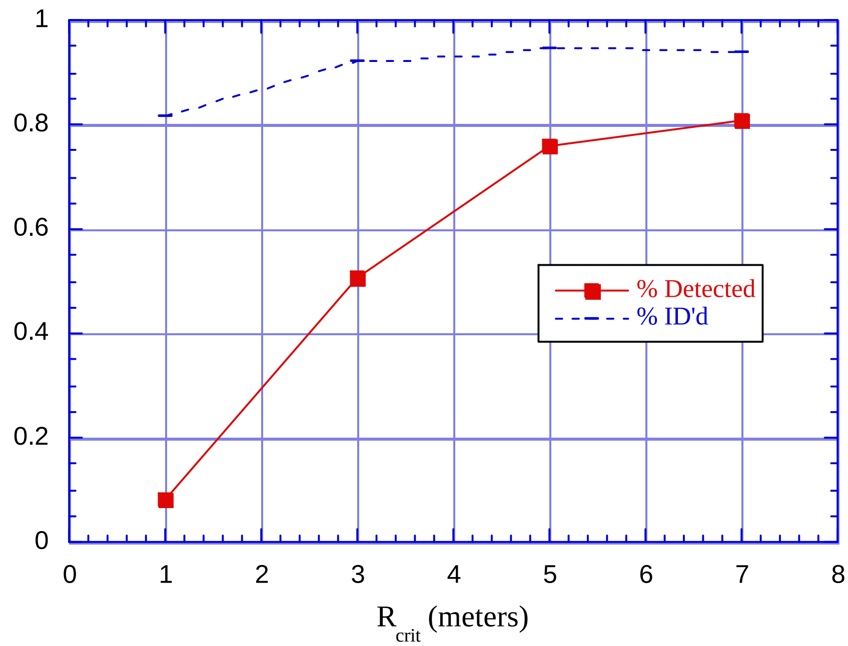
We note that our calculation of P_d does not account for holiday areas, as described in Section II.C above, and thus our value is likely to be a lower bound estimate of the capability of this system. From the 'raw' data provided to us, we had no way of determining which targets were not overflown by the REMIDS system, but it is likely that this is only a few percent effect.[11]

[10] The effect of the offset bias clearly increases as R_{crit} decreases: for example, at R_{crit} = 3m, the offset bias correction would yield a P_d of 56%, an 11% improvement over the uncorrected value of 51% shown in Fig. 3.2.

[11] The REMIDS team estimates that the overflown area was about 96% of the total area. If we assume a random distribution of targets, then we might reasonably expect a P_d of 0.76/0.96 = 79%, relative to the overflown area.

To estimate the false alarm rate, we looked at three regions of the site that contained no targets. The area of the first region was 0.073 km^2, and the REMIDS system declared a total of three targets in that area, identifying all three as gator mines. This translates into a FAR of 41 per km^2. The second region was 0.049 km^2 in area, and the REMIDS system declared two targets, one a gator mine, and the other UXO scrap. This also yields a FAR of 41 per km^2. The third region was a much smaller area of only 0.028 km^2, and in this region, the REMIDS system declared no targets. We conclude, therefore, that the REMIDS system, with its three-stage analysis, yields a very low false alarm rate. We emphasize, however, that the false alarm rate would have been much higher if only the first two analysis stages had been employed (spectral and size/shape), because the operator played a crucial role in recognizing and eliminating false alarms.

Fig. 3.2: REMIDS Performance at Yuma



In addition to its potential use in identifying individual UXO targets, REMIDS also has potential as a large-area search system, to determine which sections of a large area must be thoroughly investigated, which sections will require only minimal investigation at isolated spots, and which sections can be declared ordnance-free.
In Table 3., we show the number of false alarms after stage 1 and stage 2 are completed for different backgrounds in other locations at Yuma and Ft. Rucker, which were all ordnance-free. These sites represent a variety of backgrounds: a Yuma sandy/slit area that is clear of vegetation named the Drop Zone (this area is the practice drop zone for parachute training); a Yuma sandy/slit area with small foliage named Indian Wash (contains desert varnish); a Yuma area that is a bank on Mitri Lake with dense vegetation; a Yuma area called the Pavement Dyno that is a pavement-covered area; a Yuma area called Sand Dyno that corresponds to a very loose sand area

(this sand is not indigenous to the area); a Ft. Rucker slit area that is a newly plowed field; a Ft. Rucker area that is covered with new crop vegetation; and a Ft. Rucker area that has field vegetation shorter than two inches. The table is generated for D_{thresh} of 1.6, 1.0, and 0.7. The target types are plastic (Pl), white paint (WP), aluminum (Al), and oxidized iron and olive drab print (FeOD).

Table 3. Background False Alarm Rates per Km^2 from Stage 1 and Stage 2 Processing

D_{thresh}	1.6				1.0				0.7			
Target Type	Pl	WP	Al	FeOD	Pl	WP	Al	FeOD	Pl	WP	Al	FeOD
background type												
Drop Zone	500	0	0	0	500	0	0	0	0	0	0	0
Indian Wash	0	0	0	4500	0	0	0	500	0	0	0	0
Mitri Lake	20000	0	0	0	10000	0	0	0	2500	0	0	0
Pavement Dyno	0	1000	0	14000	0	1000	0	7000	0	0	0	1500
Sand Dyno	0	0	0	0	0	0	0	0	0	0	0	0
Plowed Field	0	0	0	0	0	0	0	0	0	0	0	0
New Crop	0	0	0	0	0	0	0	0	0	0	0	0
Short Vegetation	0	0	0	0	0	0	0	0	0	0	0	0

For the most part, the table indicates that the REMIDS algorithm would indeed predict much of these sites to be ordnance-free. The Drop Zone background would need further analysis (Stage 3 Operator Discrimination) for the plastic class. The Indian Wash background would need further analysis (Stage 3 Operator Discrimination) for the oxidized iron and olive drab paint class. The Mitri Lake background would need further analysis for the plastic class. Finally, the Pavement Dyno background false alarms in the white paint class were caused by the white striping on the pavement and the oxidized iron and the false alarms in the olive drab paint class would need further analysis.

CONCLUSIONS

Based on the above results, it is clear that the performance of this system is site- dependent. If the background is grassy as at Ft. Rucker, it is quite likely that a P_d of at least 90% is achievable with very low false alarm rates using only the Stage 1 analysis. At sites such as Yuma, such performance is probably impossible with just the spectral information. However, fair-to-good performance can be expected once the size/shape filter and operator filter are employed, *if* one is willing to accept that the system cannot be used to detect all ordnance. Specifically, if there is a high probability of large numbers of false alarms, a minimum size requirement will probably be required, and the system will therefore not be able to detect items smaller than this minimum size. Thus, if one were to utilize REMIDS to detect surface UXO larger than four pixels at a site such as Yuma, then one might be able to expect a P_d of about 76%, assuming a declaration radius of 5 m, with a very low false alarm rate. (We note again that this number is probably a lower-bound estimate of P_d, because holiday areas were not accounted for in the calculation.)

On the other hand, in realistic scenarios with large areas, the methodology adopted at Yuma would have to be modified, because too much emphasis was placed on the ability of the operator to discriminate the ordnance items. This modification would entail a more sophisticated shape/size filter to pre-screen the objects, so that the operator could complete his role in a reasonable amount of time. It is not known at this time whether transferring the bulk of the discrimination from the operator to the size/shape filter will result in similar performance as that achieved at Yuma, although it is likely that at least some degradation in performance may be expected.

Although we cannot state for certain whether this system can be relied upon to predict individual ordnance types with high P_d and low P_{fa} over large areas with potentially challenging backgrounds, we do feel reasonably confident in claiming that this system has potential in a large-area search mode of operation, with minimal operator participation.[12] It is often the case that when large tracts of land must be cleared of UXO, a significant portion of that land is already clean, but it is not known which portions are clean and which must be cleared. From the previous discussion, it is evident that the REMIDS system could play a valuable role in clearing large tracts of land by determining quickly which areas are already likely to be clean, and which warrant further investigation. This analysis can be performed solely with the spectral filter and the crude size filter employed at Yuma. More sophisticated size/shape filters will render this system even more effective in screening large areas. In any case, because the spectral and shape/size filters do not require much operator input, this screening can be performed very quickly and reliably.

REFERENCE

Ballard, John H., R. M. Castellane, B. H. Miles, K. G. Wesolowicz. (1992) "The Remote Minefield Detection System (REMIDS) II Major Components and Operation." Vicksburg, MS: US Army Engineer Waterways Experiment Station.

[12] The only way in which operator skill would be introduced in this application would be in determining the 3N+1 parameters for the spectral filter.