Statistical sensor fusion analysis of near-IR polarimetric and thermal imagery for the detection of mine-like targets

Robert A. Weisenseel\textsuperscript{a} * W. Clem Karl\textsuperscript{a} David A. Castañón\textsuperscript{a}
Charles A. DiMarzio\textsuperscript{b}

\textsuperscript{a}Boston University, Electrical and Computer Engineering Department
\textsuperscript{b}Northeastern University, Electrical and Computer Engineering Department

\section*{ABSTRACT}

We present an analysis of statistical model based data-level fusion for near-IR polarimetric and thermal data, particularly for the detection of mines and mine-like targets. Typical detection-level data fusion methods, approaches that fuse detections from individual sensors rather than fusing at the level of the raw data, do not account rationally for the relative reliability of different sensors, nor the redundancy often inherent in multiple sensors. Representative examples of such detection-level techniques include logical AND/OR operations on detections from individual sensors and majority vote methods. In this work, we exploit a statistical data model for the detection of mines and mine-like targets to compare and fuse multiple sensor channels.

Our purpose is to quantify the amount of knowledge that each polarimetric or thermal channel supplies to the detection process. With this information, we can make reasonable decisions about the usefulness of each channel. We can use this information to improve the detection process, or we can use it to reduce the number of required channels.

Keywords: mine, sensor fusion, statistical, multispectral, polarimetric, thermal, REMIDS, UXO, detection

\section{1. INTRODUCTION}

The need for mine and unexploded ordnance (UXO) detection and removal is growing in both military and humanitarian applications. In places like Cambodia, the threat of mines to the general populace is overwhelming. Since the end of the Cold War, there has been a growth in smaller, regional conflicts where threats are often not from high-tech weaponry, but from inexpensive ordnance. Bosnia is just one example. Mines are one of the least expensive weapons available, and their threat often far outlasts the conflict for which they are emplaced. Mine detection is therefore necessary in two roles: detecting minefield obstacles for military intelligence and detecting individual mines for eventual removal.

One specific subject of mine detection involves wide-area surveillance. In one case, military forces need to chart possible impediments to ground movement accurately over broad swaths of territory. In humanitarian applications, surveyors examining terrain for mine clearance can limit the area searched with wide-area surveillance. For both situations, it is highly desirable to conduct mine searches from the air to minimize the danger to personnel. Unfortunately, ground penetrating radar and quasistatic electromagnetic approaches are somewhat limited in range.

Optical techniques can meet many of the requirements of wide-area minefield detection. One area of much interest is polarimetric sensing. In the near infrared domain, surface mines have a highly polarizing characteristic. Disturbed soil, indicative of buried objects, may also have such a polarizing nature. In fact, mine detection rates are high using only polarization information, but reducing the false alarm rate is a harder problem, requiring that we apply additional information. While this feature alone is inadequate to detect mines, it can be a powerful tool when combined with multispectral information, such as thermal imaging or imaging spectrometers.

However, fusing information from multiple sources in a rational, statistical way is often neither simple nor obvious. In previous work, we showed that fusing local spatial information with polarization and thermal observations can reduce false alarms, either through a statistical prior model or local adaptivity.\textsuperscript{1} In this paper, we focus solely on the

\textsuperscript{R.A.W.: Email:lorax@bu.edu; WWW: http://mdsp.bu.edu/lorax.html}
relative advantages of fusing combinations of polarization, reflectance, and thermal data for distinguishing mine-like targets from background. Counterintuitively, we find that fusing only polarization and thermal information gives better results than any other combination, including the combination of all channels together, even under conditions of low thermal contrast. This finding is likely caused by model mismatch, but it clearly points out the complexity of choosing an appropriate fusion method. System designers should not develop sensors, models, and algorithms in isolation. We must develop these technologies together to achieve optimal solutions.

2. SENSOR BACKGROUND

Our data consist of three infrared imagery channels generated by the U.S. Army Engineer Waterways Experiment Station’s Remote Minefield Detection System (REMIDS).\textsuperscript{2,3} The system images the first two channels using an active infrared system at the $1.05 \, \mu m$ wavelength. One channel is percent polarization, $(P - S)/(P + S)$, and the other channel is total reflectance, $(P + S)$, where $P$ is reflectance in parallel polarization and $S$ is reflectance in cross polarization. The third channel is a passive thermal infrared channel operating over the 8-12 $\mu$m range. The sensor is mounted on an airborne platform, represented in Figure 1. The data that we use for this study has a resolution of 2-3 inches per pixel, depending on altitude, although later system upgrades have increased this resolution.

Several researchers have shown that polarization characteristics are advantageous in the mine detection process.\textsuperscript{2,4,5} Man-made objects such as mines and other unexploded ordnance (UXO) are significant near infrared polarizers compared to natural backgrounds, which tend to be uniformly random polarizers. Disturbed earth can indicate buried objects, and may also be a source of polarization features. REMIDS only uses linear polarization features, but we can extend the same algorithmic techniques to circular polarization as well. Additionally, thermal data alone is sensitive to weather and time of day, especially diurnal thermal crossover.

Figure 1. REMIDS Thematic Representation
The active near-IR sensor used in REMIDS is capable of all-weather imaging, but this all-weather capability comes at the price of a laser source, its power supply, and other support equipment. A passive polarization sensor could be implemented cheaply, but would have weather-dependent detection and false alarm rates. An alternative system is a passive hyperspectral polarimetric imager, but this again increases the system cost. Additionally, the computational requirements of the system increase approximately as the square of the number of channels used in the detection process, so every channel that we can ignore reduces the overall cost of the system. This last cost is the main reason we are fusing different combinations of sensors in this study: to determine if intelligently choosing channels to fuse can provide such savings with little loss in detection capability.

One difficulty with the data set used in this study is that the percent polarization channel had been saturated prior to the signal processing stage to “improve” usage of available dynamic range. Unfortunately, doing so results both in a decrease in signal-to-clutter ratio in the percent polarization channel and an unrealistic decrease in the variance of the mines in this channel. This approach also saturated many background pixels, and thus these pixels cannot be distinguished from mine pixels through the percent polarization channel alone. To distinguish these pixels, we must use data from the other channels or prior information. We will return to this issue in Section 4.

### 3. METHOD

To analyze and compare the advantages of the various channels and their combinations, we choose to use a simple Maximum Likelihood (ML) approach to detection:

\[
\hat{x} = \arg \max_x p(y|x) = \arg \max_x \ln p(y|x)
\]

where \( y \) is the multichannel image and \( x \) is the hypothesis image of mine locations, each represented as a vector,

\[
y = [y_1^T, \ldots, y_N^T]^T \quad x = [x_1, \ldots, x_N]^T
\]

\[
y_i = \begin{bmatrix} y_i^p \\ y_i^r \\ y_i^t \end{bmatrix} \in \mathbb{R}^3 \quad x_i = \begin{cases} 1 & \text{if pixel } i \text{ is a mine} \\ 0 & \text{if pixel } i \text{ is not a mine} \end{cases}
\]

and \( N \) is the total number of pixels. Thus, each element of the observation image, \( y_i \), is a 3-vector where the superscripts \( p, r, \) and \( t \) denote the three channels, percent Polarization, total Reflectance, and Thermal, respectively. For our channel comparison analysis, we sometimes ignore one or two channels. For these cases, \( y_i \) becomes either a 2-vector or a scalar, respectively.

For our observation model, we assume that each pixel of data, \( y_i \), is conditionally independent of all other data pixels, conditioned on knowledge of mine presence at that pixel, \( x_i \). Formally,

\[
p(y|x) = p(y_1, \ldots, y_N|x_1, \ldots, x_N) = \prod_{i=1}^{N} p(y_i|x_i)
\]

This assumption allows us to perform the maximization in Equation 1 individually for each pixel,

\[
\hat{x} = \arg \max_x \ln p(y|x) = \arg \max_x \sum_{i=1}^{N} \ln p(y_i|x_i)
\]

\[
\hat{x}_i = \arg \max_{\hat{x}_i} \ln p(y_i|x_i)
\]

Thus, we formulate the problem as a likelihood ratio test,

\[
\frac{p(y_i|x_i = 1)}{p(y_i|x_i = 0)} \overset{\hat{x}_i = 1}{\geq} \alpha
\]

where \( \alpha \) is a threshold chosen by the user.
We also assume that each pixel is a conditionally identically distributed Gaussian, with parameters that are a function of the presence at the pixel of interest,

\[
p(y_i|x_i) = \frac{1}{|2\pi\Sigma(x_i)|} \exp \left\{ -\frac{1}{2} (y_i - \mu(x_i))^T \Sigma(x_i)^{-1} (y_i - \mu(x_i)) \right\}
\]

where \(\mu(x_i)\) and \(\Sigma(x_i)\) are the mean and the covariance, respectively, of the observation variable, \(y_i\). For the purposes of this paper, we assume that the mean and covariance are known for both background and mines, and we estimate the mean and a full covariance matrix for both mines and background directly from the data. We make no attempt to model or adapt to nonstationarity of the background in this study. The likelihood ratio test then simplifies as follows,

\[
[y_i - \mu(0)]^T \Sigma^{-1}(0) [y_i - \mu(0)] - [y_i - \mu(1)]^T \Sigma^{-1}(1) [y_i - \mu(1)] \sim \chi^2 \gamma
\]

Note that these decision threshold surfaces are quadratic, not linear, because the covariances are not equal.

The simplicity of this model is desirable because it makes the resulting analysis and comparison of the channels easier to understand. One possible drawback, however, is that this very simplicity may not be a good match to reality, and may produce unexpected results.

4. RESULTS

For our test we have access to only two data sets, both from 1991 test flights of REMIDS over Fort Drum, New York. The first flight took place at 9 am on an overcast day, resulting in low thermal contrast between mines and background, but also low thermal variance. The second flight occurred at 3 pm on a clear, sunny day. We show representative segments of all channels of both scenes in Figures 2 and 3. The mines in these images are surface patterned anti-tank mines, but this approach to comparing the relative usefulness of the channels is equally applicable to scattered mines as long as sufficient spatial resolution is available. The applicability to buried mines depends on whether the advantages of polarization information extend to buried mines that disturb the surface of the soil in which they lie. Thermal detection of buried mines is an established technique.

For the purposes of our analysis, we define a detection as any continuously 8-connected region of pixels labeled as a mine by the ML approach. Using an image ground-truthed by hand, we define a correct detection as any single detection that coincides with at least one pixel of any continuously 8-connected region of pixels in the ground-truth image. If two such detection regions coincide with a mine region, we only count one as a correct detection. We define the total number of false alarms as the difference between the total number of detections and the number of correct detections. This definition is imperfect; as the threshold, \(\gamma\), of Equation 6 increases, an initially continuous detection region can break into multiple disjoint regions, resulting in an increase in the number of false alarms, without a related increase in correct detections. Using a Maximum A Posteriori (MAP) detector with a smoothness prior model can relieve this problem.

We show the resulting empirically derived Receiver Operating Characteristic (ROC) curves for the clear and cloudy days in Figures 4 and 5, respectively. Each graph contains multiple curves, one for each possible combination of the three channels: percent polarization, total reflectance, and thermal. We label the graph legends with the same convention used in the superscripts of Section 3: ‘p’ for percent Polarization, ‘r’ for total Reflectance, and ‘t’ for Thermal. We denote combinations of channels by multiple letters, e.g. ‘rt’ represents the combination of the total reflectance and thermal channels. Since these are semilog plots, we assigned a value of 0.99 to any cases of zero false alarms, for display purposes.

There are three unusual features of these ROC curves that we must interpret. We explained the first such feature previously, the occasional increase in false alarms as correct detections decrease, as the breaking of a detection into multiple smaller detections as detection threshold increases. The second is the square ROC curve associated in each graph with the polarization channel alone: a sudden decrease from 1 to 0 in the probability of correct detection as you trace the curve from the right with no change in false alarms, followed by an immediate decrease in the number of false alarms to zero. This odd characteristic is a result of the saturation of the polarization channel described in Section 2. There are a fixed set of false alarms and correct detections that exactly attain the upper limit of the range.
of values on which the polarization channel is defined. As the detection threshold approaches this upper limit it rules out all other false alarms. As the threshold crosses the top of the range, both correct detections and false alarms plummet to zero immediately. While the graph we chose to represent this is not unique, it is the closest match to our intuition about the shape and structure of ROC curves.

The third unusual feature involves the order of the ROC curves from lower right to upper left. From a statistical viewpoint, adding channels, and hence adding information, cannot result in a worse detection characteristic unless the model used does not match the data. Yet our ROC curves in both types of weather clearly show an increase in false alarms as the reflectance channel is included with the polarization and thermal channels. Similarly, the reflectance channel alone appears to perform better than the combined thermal and reflectance channels on the overcast data. There are two possible ways that we can explain this counterintuitive result. One is our simplifying assumption of background stationarity, which a cursory perusal of the data will show is clearly inaccurate. Including model adaptivity may fix this problem. Another possibility could be a sensitivity in the reflectance channel to the parameter estimates, particularly its covariance with the thermal channel. Each of the unusual cases involves combining the reflectance channel with the thermal channel.

In the other cases, the ROC curves follow our intuition about fusing channels: fusing one channel with another results in an improvement over either channel separately. The interesting result is the amount of improvement. First, note that the polarization and reflectance false alarm rates, alone and separately, do not change much as a function of the weather conditions. We could anticipate this, since a narrowband active sensor should be influenced little by ambient radiation. The thermal channel and all of its combinations outstandingly improve on a clear, sunny day. Again, we could anticipate this since thermal contrast increases dramatically with input radiation. Surprisingly, even on the overcast day, the thermal channel in conjunction with the polarization channel gives a significant and useful improvement, more than halving the number of false alarms of either channel separately. The reflectance channel, on the other hand, provides a much more limited improvement when used in conjunction with the polarization channel under all conditions.

5. CONCLUSIONS

In this research, we have examined the relative value of polarization, reflectance, and thermal information in the particular context of the mine detection problem. Using a particular simplified model we have shown that polarization and thermal information, when fused in a statistically rational way, can significantly improve the detection process over either alone, primarily through a reduction in false alarm rate. Also the thermal channel provides significant information, even under conditions of low thermal contrast.

The advantages of expending the computational effort to include the reflectance channel appear more dubious. The improvements in false alarm rate that this channel provides are relatively small when compared to the other two channels. Also, it appears from a cursory examination of the results that the reflectance channel may be particularly sensitive to model mismatch. However, a careful analysis of the causes of this difficulty will require further study.

6. ACKNOWLEDGMENTS

We would like to thank Ricky Goodson and Ernesto Cespedes of the US Army Corps of Engineers Waterways Experimental Station’s Environmental Laboratory for graciously sharing their REMIDS data. This work is partially supported by the Army Research Office under Grant ARO DAAG55-97-1-0013, the Air Force Office of Scientific Research under Grant F49620-96-1-0028, and the National Institutes of Health under Grant NINDS 1 R01 NS34189.
REFERENCES


Figure 2. 9 am, Overcast, Percent Polarization, Reflectance, and Thermal
Figure 3. 3 pm, Clear, Percent Polarization, Reflectance, and Thermal
**Figure 4.** Experimental ROC for ML Detector - Clear Day

**Figure 5.** Experimental ROC for ML Detector - Overcast Day