PHENOMENOLGICAL MODEL INVERSION WITH FISHER INFORMATION METRICS FOR UNEXPLODED ORDNANCE DETECTION

Jeremiah J. Remus[1] and Leslie M. Collins[2]

[1]Dept. of Electrical and Computer Engineering
Clarkson University, Potsdam, NY 13699
jremus@clarkson.edu

[2]Dept. of Electrical and Computer Engineering
Duke University, Durham, NC 27708
lcollins@ee.duke.edu

1. INTRODUCTION

Many strategies for UXO detection and discrimination utilize phenomenological models to produce the features used as inputs to statistical decision algorithms. Therefore, it is necessary to develop robust model inversion techniques to backfit the forward phenomenological model to the measured data. Model inversion techniques aim to minimize fit error between the data and the model, often using gradient descent [1] or stochastic techniques [2]. In practice, it is possible for the model inversion process to generate outliers in the parameter space, when the model inversion process returns parameter values that result in a low fit error but differ by orders of magnitudes from the parameter values for other UXO and non-UXO clutter. We wish to reduce the occurrence of these outlier anomalies, particularly for UXO. This suggests the need for a more sophisticated metric that relies on more than simple metrics such as fit error or L-norm.

The standard model fit error measures do not incorporate the spatial distribution of the data used in the model inversion. Recent studies have focused on data quality as an important component in the model inversion process and have proposed different figures of merit for quantifying data quality (e.g. [3, 4]). These studies suggest that the quality of the data can be determined based on particular data collection characteristics such as measurement spatial density and signal-to-noise ratio. This study presents a data-dependent evaluation of the model inversion outcomes that uses the Fisher information [5] on the model parameters. Previous data quality metrics have relied primarily on data density and estimates of SNR, and do not directly consider the specific phenomenological model. This study incorporates the Fisher information in a joint metric optimization to assess the spatial distribution of data and how well the model parameters are supported by the data used in the model inversion.

2. MODEL INVERSION USING A JOINT METRIC OPTIMIZATION

The model used in this study was a nonparametric extension of the standard dipole model of Carin et al. [6]. In the nonparametric model, the magnetization tensor matrix M is recalculated at each time gate, such that the values on the diagonal are functions of time:

$$M = \begin{bmatrix} k_2(t) & 0 & 0 \\ 0 & k_2(t) & 0 \\ 0 & 0 & k_1(t) \end{bmatrix}.$$

In addition to the k_i parameters, there are five extrinsic model parameters that define the object's rotation {θ, Φ} and location {dx,dy,dz} in space. The standard, baseline model inversion process is based on gradient descent procedures that minimize an error norm (typically L^2) between the measured data and the model predictions. Due to the presence of many local minima, the final value is dependent on the initial values of the parameters used to start the model inversion. To improve the likelihood of finding the global best-fit parameters, the model inversion is restarted with different initializations of the parameters. The best-fit parameter values are selected from the set

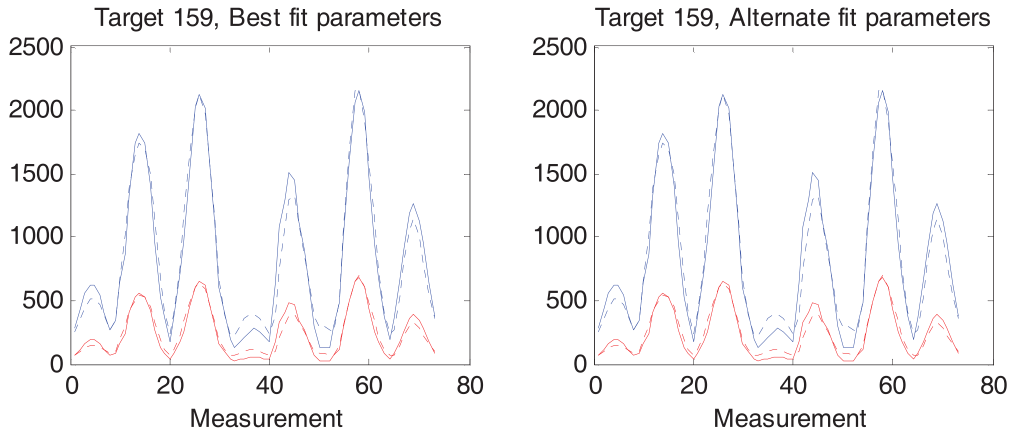

Figure 1. Plots of the measured data (solid lines) and model fits (dashed lines) for the two optimization criteria. The blue lines correspond to the first time gate and the red lines correspond to the third time gate. The left subplot used fit error only in the model inversion; the right subplot used the joint metric optimization.

of restarts according to some metric (often the same squared error measure optimized in the gradient descent procedure).

The parameters k_i are of primary interest since they will be used to generate the features used for UXO classification. Thus, in the new approach, just these parameters were included in the Fisher information matrix while the five extrinsic model parameters were treated as nuisance parameters and excluded from the information metric. In the calculation of the Fisher information metric, the sensor measurements are assumed to be of the form $\mathrm{Y} = f(\beta_{INT}, \beta_{EXT}) + Z$, where $f(\beta_{INT}, \beta_{EXT})$ is the dipole model response and Z is white Gaussian noise. The dipole model response is a function of the set of intrinsic model parameters $\beta_{\mathrm{INT}} = \{k_1, k_2\}$ and the set of extrinsic model parameters $\beta_{\mathrm{EXT}} = \{\theta, \Phi, dx, dy, dz\}$. The Fisher information matrix takes the form $J = \sum_{i=1}^{N} J_n$, where $J_n = [\nabla f(\beta_{INT}; \beta_{EXT})][\nabla f(\beta_{INT};\ \beta_{EXT})]^{\mathrm{T}}$ and J_n is evaluated at the n^{th} measurement location.

The Fisher information metric alone is not a sufficient optimization criterion for model inversion, due to the assumption in the calculation of the Fisher information matrix that the sensors measurements are based on the model with additive white Gaussian noise. Thus, for a given set of model parameters, the Fisher information metric is maximized for a particular pattern of measurement locations; the actual measurements are not used in the calculation. Instead, the Fisher information was included when selecting the best fitting parameters from the set of parameters generated by all the restarts of the optimization routine. Typically, the restarts are sorted by fit error, and the parameters corresponding to the lowest fit error are selected. In the joint metric optimization, the restarts are given two scores: the rank when sorted by squared error, and the rank when sorted by Fisher information. In the joint metric optimization, the model parameter set with the minimum sum of the two ranks is selected from the set of restarts.

3. RESULTS

The two model inversion methods were tested on a data set from Camp Sibert, Alabama that consisted of 175 anomalies: 59 UXO and 116 non-UXO clutter. The data was collected using a Geonics EM-61 MkII towed sensor array. The measured data were modeled using the nonparametric dipole model described above. In the standard inversion process, an iterative trust-region method of gradient descent was used to find the model parameters that result in the best fit (L^2 norm) to the measured data. The gradient descent was restarted 500 times, each time using a randomly-generated initialization of the model parameters. In the standard model inversion process, the parameter set that resulted in the lowest overall squared error out of the 500 restarts was selected as the best fitting parameters. In the joint fit error / Fisher information model inversion process, the parameter set that minimized the joint metric was selected from the 500 restarts as the best fitting model parameters. The model parameters were used to calculate a set of features for classification as were processed in Lhomme et al [4]. The

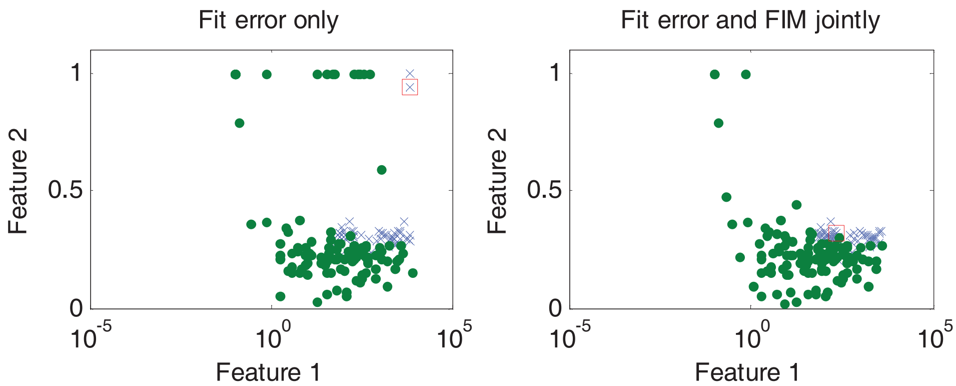

Figure 2. Plots of feature space that result when using fit error only as the optimization metric (left subplot) or when using the joint metric optimization (right subplot). The UXO are represented by x symbols; the non-UXO clutter are represented by · symbols. The square identifies the UXO example shown in Figure 1. Note the lack of outlier UXO in the right subplot when using the joint metric optimization.

objects were represented as a two-dimensional feature vector calculated as $\left[k_1(1), \max\left\{\frac{k_2(3)}{k_2(1)}, \frac{k_1(3)}{k_1(1)}\right\}\right]$. The first feature corresponds to the magnetization tensor component fit to the major axis at the first time gate (t = 1). The second feature is related to the decay rate and is calculated as a ratio between the magnetization tensor components for the first and third time gates.

The benefit of adding the Fisher information metric to the model inversion process is first shown using a single target as an example. The target of interest is a 4.25" mortar at a depth of 37 cm. The results of the model fitting are shown in Figure 1. The left subplot corresponds to the standard model inversion; the right subplot corresponds to the joint metric optimization model inversion. Visually, both sets of parameter values produce model outputs that appear to fit the measured data with similar degrees of error. The difference in the sum of squared errors is less than 2.5%. However, the set of parameters producing the model fit shown in the right subplot has a Fisher information metric that is two orders of magnitude larger than the parameters that produce the model fit shown in the left subplot. This indicates that the data collection for this target provides much stronger support for these model parameters. These parameter values are also more consistent with the majority of the UXO targets (shown in Figure 2), which allows for improved classification performance. In Figure 2, each subplot shows the two-dimensional feature space; the x symbols represent UXO and the · symbols represent non-UXO clutter. The left subplot shows features generated from model parameters found using fit error only; the right subplot used the joint metric optimization. A square symbol identifies the target plotted in the Figure 1; it can be seen that using the standard fit-error-only optimization results in this target appear as an outlier. However, using the joint metric optimization, the occurrence of outliers is reduced. Classification performance is also improved due to the reduction of UXO outliers. Leave-one-out (LOO) training and testing was used with the Distance Likelihood Ratio Test (DLRT) classifier [7] to compare the two feature sets. Figure 3 shows ROC curves produced using the two feature sets (standard fit-error-only and the joint metric optimization). The feature set generated using the joint metric optimization results in substantially better ROC performance at high P_D levels and reaches the desired P_D = 100% level as a significantly lower P_{FA} than when using the features generated by the standard model inversion process.

4. CONCLUSIONS

Phenomenological models are a component in many of the current strategies for UXO detection, requiring the use of numerical optimization techniques for model inversion. This study considered a joint metric optimization that added a Fisher information metric to the standard model fit error metric. The Fisher information metric provides a measure of how well the model parameter values are supported by the measurements used in the model inversion. Adding the Fisher information metric to the model inversion process reduces the occurrence of anomalous model parameters, resulting in a more robust model inversion process.

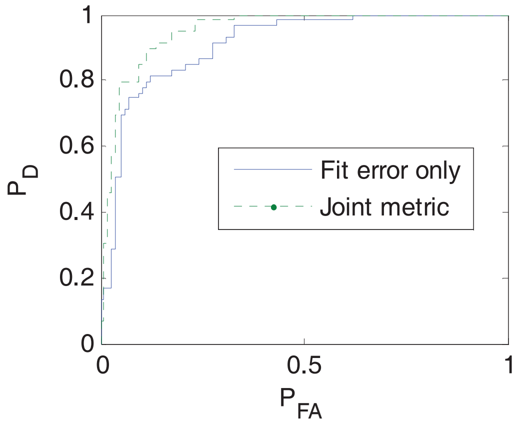

Figure 3. Comparison of classification performance when using the two feature sets generated from the different optimization metrics.

References

[1] K. Levenberg, "A method for the solution of certain non-linear problems in the least squares," *Quarterly of Applied Mathematics,* vol. 2, pp. 164-168, 1944.
[2] J. Stalnaker and E. Miller, "Particle swarm optimization as an inversion tool for a nonlinear UXO model," in *International Geoscience and Remote Sensing Symposium,* 2007, pp. 432.
[3] S. E. Walker, L. R. Pasion, D. W. Oldenburg and S. D. Billings, "Investigating the effect of data quality on time domain electromagnetic discrimination," *Journal of Applied Geophysics,* vol. 61, pp. 254-278, 2007.
[4] N. Lhomme, D. W. Oldenburg, L. R. Pasion, D. B. Sinex and S. D. Billings, "Assessing The Quality of Electromagnetic Data for The Discrimination of UXO Using Figures of Merit," *Journal of Environmental and Engineering Geophysics,* vol. 13, pp. 165-176, 2008.
[5] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* New York: Wiley-Interscience, 1991,
[6] L. Carin, H. Yu, Y. Dalichaouch, A. R. Perry, P. V. Czipott and C. E. Baum, "On the Wideband EMI Response of a Rotationally Symmetric Permeable and Conducting Target," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 39, pp. 1206-1213, 2001.
[7] J. J. Remus, K. D. Morton, P. A. Torrione, S. L. Tantum and L. M. Collins, "Comparison of a distance-based likelihood ratio test and k-nearest neighbor classification methods," in *IEEE Workshop on Machine Learning for Signal Processing,* 2008, pp. 362-367.