

# Statistically-Based Methods for Anomaly Characterization in Images from Observations of Scattered Radiation\*

Eric L. Miller  
235 Forsyth Building  
Northeastern University  
360 Huntington Ave.  
Boston, MA 02215  
Tel: (617) 373-8386  
email: elmiller@cdsp.neu.edu

February 14, 1997

## Abstract

In this paper we present an algorithm for the detection, localization, and characterization of anomalous structures in an overall region of interest given observations of scattered electromagnetic fields obtained along the boundary of the region. Such anomaly detection problems are encountered in applications including medical imaging, radar signal processing, and geophysical exploration. The techniques developed in this work are based on a non-linear scattering model relating the anomalous structures to the observed data. A sequence of M-ary hypothesis test are employed first to localize anomalous behavior to large areas and then to refine these initial estimates to better characterize the true target structures. We introduce a method for the incorporation of prior information into the processing which reflects constraints relevant directly to the anomaly detection problem such as the number, shapes, and sizes of anomalies present in the region. The algorithm is demonstrated using a low-frequency, inverse conductivity problem found in geophysical applications.

Submitted to *IEEE Trans. on Image Processing*  
EDICS: IP 2.5 Geophysical and Seismic Imaging

---

\*This work was supported in part by the Department of Energy under grant DOE-FC07-95ID13395 and the Center for Electromagnetics Research at Northeastern University

# 1 Introduction

A common problem arising in a variety of application areas is the non-invasive detection, localization, and characterization of anomalous areas in an overall region of interest given observations of scattered acoustic or electromagnetic radiation collected along a portion of the region's boundary [9, 15]. For example, tumor detection in otherwise healthy tissue arises frequently in many areas of medical signal processing [1]. In geophysical exploration, particularly low values of electrical conductivity in a region of the earth provide one indication as to the presence of oil [18]. Finally, there has been significant interest recently in the use of ground penetrating radar for the detection of buried objects such as land mines and metallic drums containing hazardous waste [14].

Typical methods for solving this *anomaly detection* problem proceed by first forming an image of distribution of a particular physical quantity (eg. sound speed or electrical permittivity) internal to the medium and subsequently post-processing the reconstructed profile to locate and classify anomalous areas [9, 15]. The problem of image formation from scattered field measurements (also known as the inverse scattering problem) presents a collection of difficulties which makes this image-then-detect approach rather unattractive. First, because the reconstructions are obtained typically over a fine scale grid of pixels, these problems are computationally intensive. The complexity of many practically interesting problems is further increased by the non-linear relationship between the physical parameter of interest and the observed scattered fields. The resulting imaging algorithms require the solution of a large-scale, non-convex, non-linear optimization problem [18].

Additional problems arise from the fact that one often is restricted to probing the medium with sources operating at only a few frequencies and observing the scattered fields over a limited portion of the region's periphery. As discussed in [11, 13] under these circumstances the imaging problem is highly ill-posed in that the information content in the the data does not support a stable fine-scale reconstruction of a pixelated representation of the physical parameter. Commonly, this problem

of ill-posedness is solved through the use of a regularization scheme which enforces a degree of smoothness in the reconstruction [10]. The use of such a regularizer for anomaly characterization serves to *increase* the difficulty of detection and identification after image formation. Typical targets represent regions where the medium's structure is locally *not* smooth. Imaging with a smoothness regularizer results in a low-pass image in which these distinguishing anomaly characteristics are softened thereby increasing the difficulty of subsequent detection and identification.

Over the past decade, significant research has been performed on the detection of anomalies directly from scattered field data. Rather than solving a large, ill-posed imaging problem, these techniques employ statistical signal processing methods to extract from the data a relatively small number of geometric parameters describing the distribution of anomalous structures. In [15], Rossi and Willsky considered the use of estimation-theoretic analysis and algorithmic methods for determining the location of a single object of known size and structure given noisy and sparse computed tomography (CT) measurements. Devaney and co-workers [3, 17] extended these results to diffraction tomography (DT) and exact scattering applications. In [2], Bressler, Fessler and Macovski presented an algorithm for determining the locations and geometric structures of an unknown number of anomalies based on CT measurements. A similar problem is considered in [16, 20], where the authors were concerned with the detection of an unknown number of circular objects of unknown radii given noisy CT observations. The iterative algorithms in [16, 20] began by identifying a large collection of likely targets with subsequent steps (a) pruning away structures which were deemed unlikely to be actual objects (b) refining the estimates of the radii and locations for the remaining objects.

A decision-theoretic, multiscale algorithm based on a sequence of M-ary, hypothesis tests was presented and analyzed in [11] for the characterization of an unknown number of anomalous regions of unknown size, shape and amplitude for problem in which a linear model related the anomalies to

the measurements. The approach in [11] began by identifying a small number of large sized regions in which anomalies were likely to exist. Subsequent processing steps refined these estimates and provided for the re-partitioning of each coarse scale region to account for the existence of multiple structures. The application of interest was a diffraction tomography problem arising in geophysical exploration.

In this paper, the problem of detecting and localizing an unknown number of anomalous areas is addressed for problems where linear scattering models are not valid so that the full, nonlinear physics must be incorporated into the processing. As in [11], the fundamental idea behind the approach presented here is the use of a sequence of hypothesis tests which begin by localizing anomalous behavior to a few, large sized areas and subsequently “zoom in” to better localize the true target structures. It should be noted that the spatial subdivision process considered here is quite different from that in [11]. In addition to the use of a non-linear scattering model, here we introduce a decision-theoretic regularization method which allows for the inclusion of prior information specifically tailored to the anomaly detection problem. Expectations regarding the number, sizes and shapes of anomalous areas are easily and naturally incorporated into the hypothesis testing process. While we do not examine the issue extensively in this work, the anomaly detection algorithm has a structure making it well tailored for a computationally efficient, parallel implementation. As in [11], we concentrate on a low-frequency geophysical inverse conductivity problem.

In Section 2 of this paper the scattering model and problem of interest are presented. An overview of the relevant detection methods is provided in Section 3. The multiscale anomaly detection algorithm is described in Section 4 and examples provided in Section 5. Section 6 is devoted to conclusions.

## 2 Problem Formulation

### 2.1 The Scattering Model

We consider a two-dimensional inverse conductivity problem illustrated in Fig. 1 where there exist a set of electromagnetic line-sources oriented perpendicularly to the page emitting time-harmonic waves into a lossy medium. The medium is characterized by the free space values for electric permittivity,  $\epsilon_0$ , and magnetic permeability,  $\mu_0$  and a conductivity profile modeled as the sum of a constant background,  $g_0$ , and a collection of anomalies described by the function  $g(\mathbf{r})$ . The goal of the problem is to characterize  $g(\mathbf{r})$  from noisy, pointwise observations of the component of the scattered electric field oriented perpendicularly to the page. These measurements are obtained along receiver arrays positioned on the vertical edges of  $A$  from  $K$  scattering experiments. Each experiment produces a vector of measurements,  $\mathbf{y}_i$ , comprised of the in-phase and quadrature components of the scattered field obtained over a single receiver array due to energy put into the medium from one of the sources operating at a frequency,  $f_i = \omega_i/2\pi$ .

As discussed in [18], the relationship between  $g(\mathbf{r})$  and the  $j$ th element of  $\mathbf{y}_i$ ; that is, the measured scattered field at position  $\mathbf{r}_j$  is

$$y_i(\mathbf{r}_j) = \imath\omega_i\mu_0 \int_A G_i(\mathbf{r}_j, \mathbf{r}') E_i(\mathbf{r}') g(\mathbf{r}') d\mathbf{r}' + n_i(\mathbf{r}_j) \quad (1)$$

where  $n_i(\mathbf{r}_j)$  represents additive noise,  $G_i(\mathbf{r}, \mathbf{r}')$  is the Green's function for the problem and  $E_i(\mathbf{r})$  is the component of the total electric field perpendicular to the page. From [18],  $E_i(\mathbf{r})$  satisfies

$$E_i(\mathbf{r}) = \bar{E}_i(\mathbf{r}) + \imath\omega_i\mu_0 \int_A G_i(\mathbf{r}, \mathbf{r}') g(\mathbf{r}') E_i(\mathbf{r}') d\mathbf{r}' \quad (2)$$

with  $\bar{E}_i(\mathbf{r}) = I_i \imath\omega_i\mu_0 G_i(\mathbf{r}, \mathbf{r}_i)$  the background field generated by a line source with current density  $I_i$  positioned at  $\mathbf{r}_i$ . Finally  $G_i(\mathbf{r}, \mathbf{r}') = \imath/4 H_0^{(1)}(k_{i,0}|\mathbf{r} - \mathbf{r}'|)$  is the Green's function for this problem with  $k_{i,0}^2 = \omega_i^2\mu_0(\epsilon_0 + \imath g_0/\omega_i)$  and  $H_0^{(1)}(z)$  the zeroth order Hankel function of the first kind.

A discrete representations of (1) is obtained using the method of moments (MOM) [8] where  $g(\mathbf{r})$  and  $E_i(\mathbf{r})$  are expanded in pulse bases (i.e. zeroth order splines). A Galerkin scheme is employed

to discretize (2) using the same pulse bases. Upon discretization, (1) and (2) reduce to

$$\mathbf{y}_i = \mathbf{G}_{i,s} \mathcal{D}(\mathbf{E}_i) \mathbf{g} + \mathbf{n}_i \quad (3a)$$

$$\mathbf{E}_i = \bar{\mathbf{E}}_i + \mathbf{G}_i \mathcal{D}(\mathbf{g}) \mathbf{E}_i \quad (3b)$$

where  $\mathbf{g}$  is a lexicographically ordered vector of expansion coefficients for  $g(\mathbf{r})$ ,  $\mathbf{E}_i$  (resp.  $\bar{\mathbf{E}}_i$ ) is a vector of coefficients for  $E_i(\mathbf{r})$  (resp.  $\bar{E}_i(\mathbf{r})$ ),  $\mathbf{G}_{i,s}$  (resp.  $\mathbf{G}_i$ ) is a matrix representation of the integral kernel in (1) (resp. (2)), and  $\mathcal{D}(\mathbf{x})$  is the diagonal matrix whose  $(i, i)$ th element is the  $i$ th component of the vector  $\mathbf{x}$ . Solving for  $\mathbf{E}_i$  in (3b) and substituting the result in (3a) gives

$$\mathbf{y}_i = \mathbf{G}_{i,s} \mathcal{D} \left\{ [\mathbf{I} - \mathbf{G}_i \mathcal{D}(\mathbf{g})]^{-1} \bar{\mathbf{E}}_i \right\} \mathbf{g} + \mathbf{n}_i \equiv \mathbf{h}_i(\mathbf{g}) + \mathbf{n}_i. \quad (4)$$

In this paper,  $\mathbf{n}_i$  is taken to be additive, white Gaussian noise uncorrelated from one observation vector to the next. Thus,  $\mathbf{n}_i \sim \mathcal{N}(\mathbf{0}, r_i \mathbf{I})$  and  $E[\mathbf{n}_i \mathbf{n}_j^T] = r_i \mathbf{I} \delta(i - j)$  where  $\mathbf{I}$  is an appropriately sized identity matrix and  $\delta(n)$  is the Kronecker delta.<sup>1</sup> Finally, we collect the data from the  $K$  scattering experiments into a single vector,  $\mathbf{y}$  so that the overall observation model is given by

$$\mathbf{y} = \mathbf{h}(\mathbf{g}) + \mathbf{n} \quad (5)$$

where  $\mathbf{y}^T = [\mathbf{y}_1^T \ \mathbf{y}_2^T \ \dots \ \mathbf{y}_K^T]$ ,  $\mathbf{h}(\mathbf{g})$  and  $\mathbf{n}$  are defined accordingly, and  $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$  with  $\mathbf{R} = \text{diag}(r_1 \mathbf{I}, r_2 \mathbf{I}, \dots, r_K \mathbf{I})$ .

One difficulty with the use of (4) in the context of an inverse problem is that of computational complexity. A basic component of any inverse problem is the evaluation of  $\mathbf{y}_i$  for various values of  $\mathbf{g}$ . According to (4), each such evaluation requires the inversion of the large, dense matrix  $\mathbf{I} - \mathbf{G}_i \mathcal{D}(\mathbf{g})$ . In fact, for a discretization of  $A$  into an  $N_g \times N_g$  array of pixels, this matrix is of size  $N_g^2 \times N_g^2$ . To overcome this computational burden, the algorithm discussed in Section 4 is implemented using the nonlinear, extended Born approximation (EBA) [6] to the true scattering model. That is, for all computations,  $\mathbf{h}_i(\mathbf{g})$  is replaced by  $\mathbf{h}_i^{EBA}(\mathbf{g})$  where a detailed discussion of  $\mathbf{h}_i^{EBA}(\mathbf{g})$  may be found in [6, 13, 18]. Like the true scattering physics, the EBA is a nonlinear

---

<sup>1</sup>The notation  $\mathbf{x} \sim \mathcal{N}(\mathbf{m}, \mathbf{P})$  indicates that the random vector  $M\mathbf{x}$  has a Gaussian probability distribution with mean  $\mathbf{m}$  and covariance matrix  $\mathbf{P}$ .

function of the conductivity. For the geophysical application of interest here, the EBA has been shown to provide a highly accurate approximation of  $\mathbf{h}_i(\mathbf{g})$  even for conductivity profiles where the more common Born or Rytov linearized models are no longer valid [6,18]. The principle advantage of the EBA is that the computational complexity of the approximation is a small fraction of that associated with the exact scattering model [18].

## 2.2 The Anomaly Model

Because the MOM discretization procedure pixelates  $A$  into an array of rectangular cells and because the algorithm in Section 4 performs anomaly localization through a process of spatial subdivision, we are lead naturally to consider a representation for  $\mathbf{g}$  in which anomalous regions are defined to be superpositions of rectangular subsets of  $A$ . Referring to Fig. 2, the structure of the  $i$ th anomaly in  $A$  is defined by its magnitude, its size, and its location in  $A$ . Mathematically, the form for the anomalous behavior of the conductivity over the region  $A$  is

$$\mathbf{g} = \sum_{j=1}^{N_a} \mathbf{b}_j a_j = \mathbf{B}\mathbf{a}. \quad (6)$$

Here,  $N_a$  is the number of anomalous regions located in  $A$ ,  $a_j$  is a scalar defining the magnitude of each anomaly, and  $\mathbf{b}_j$  represents the discrete indicator function over the  $j$ th rectangular region. In (6), the column vector  $\mathbf{a}$  represents the collection of anomaly amplitude coefficients while  $\mathbf{B}$  is the matrix whose  $j$ th column is the lexicographically ordered form of  $\mathbf{b}_j$ . We note that the algorithm in Section 4 is not dependent upon the assumption of rectangular anomalies; rather structures with arbitrary shapes and orientations can be employed in principle through the appropriate specification of the matrix  $\mathbf{B}$ . Finally, the anomaly detection problem may now be stated in the following manner: Given the data and the noise statistics, determine the structure of the matrix  $\mathbf{B}$  and the vector  $\mathbf{a}$ , i.e. find the number, locations, sizes, and the amplitudes of the anomalous regions in  $A$ .

### 3 M-Ary Hypothesis Testing

The anomaly detection algorithm to be presented and analyzed in Sections 4 and 5 is based on a sequence of M-ary hypothesis tests [19] in which each hypothesis corresponds to a particular anomaly distribution in  $A$ . That is, the  $m$ th hypothesis for one such test takes the form

$$H_m : \quad \mathbf{g} = \mathbf{B}_m \mathbf{a}_m \quad (7)$$

for  $m = 0, 1, \dots, M - 1$ . In (7)  $\mathbf{B}_m$  describes the geometric characteristics of the hypothesized anomaly configuration and  $\mathbf{a}_m$  is the associated amplitude vector.

Given a data vector, the hypothesis test itself may be viewed as a rule for selecting the “best” anomaly configuration from among the  $M$  possible choices. In this paper, the following maximum *a posteriori* (MAP) procedure is employed

$$\text{Choose } H_m \text{ with } m = \arg \max_j \log [p_{\mathbf{y}|H_j}(\mathbf{Y}|H_j)] + \log(p_j) \quad (8)$$

where  $p_{\mathbf{y}|H_j}(\mathbf{Y}|H_j)$  is the probability density function for the data under  $H_j$  and  $p_j$  represents the prior probability that  $H_j$  is in fact true. The Gaussian model for  $\mathbf{n}$  allows (8) to be written as

$$\begin{aligned} \text{Choose } H_m \text{ with } m &= \arg \min_j \frac{1}{2} \|\mathbf{y} - \mathbf{h}(\mathbf{B}_j \mathbf{a}_j)\|_{\mathbf{R}^{-1}}^2 + \pi_j \\ &\equiv \arg \min_j \ell_j(\mathbf{y}) + \pi_j \end{aligned} \quad (9)$$

where  $\|\mathbf{x}\|_{\mathbf{A}} \equiv \mathbf{x}^T \mathbf{A} \mathbf{x}$  and  $\pi_j = -\log(p_j)$ .

For the anomaly detection algorithm in Section 4, the amplitude vectors,  $\mathbf{a}_m$ , are assumed to be deterministic, but unknown quantities. Thus, a generalized hypothesis test is employed in which  $\mathbf{a}_m$  in (7) is replaced by the following estimate of  $\mathbf{a}_m$  given the data

$$\hat{\mathbf{a}}_m = \arg \min_{\mathbf{a}_m > -g_0} \frac{1}{2} \|\mathbf{y} - \mathbf{h}(\mathbf{B}_m \mathbf{a}_m)\|_{\mathbf{R}^{-1}}^2 + \alpha^2 \|\mathbf{a}_m\|. \quad (10)$$

The notation  $\mathbf{a}_m > -g_0$  implies that each element of the vector  $\mathbf{a}_m$  is greater than the negative of the background conductivity and is required to ensure that the overall conductivity in region  $A$ , is greater than zero [18]. As in [11], (10) represents a regularized, maximum likelihood estimate of the amplitude vector where the second term is used to offset any ill-posedness in the estimation problem.



The solution to this constrained non-linear, least squares optimization problem is obtained using a Gauss-Newton algorithm [4] modified to account for the bound constraint in  $\mathbf{a}_m$ .

The evaluation of the each likelihood statistic,  $\ell_m(\mathbf{y})$ , requires the solution of a non-linear optimization problem; however, the dimension of each problem is equal to the number of anomalies associated with  $H_m$  which is typically far less than the number degrees of freedom found in imaging application [13, 18]. Thus, the computational burden of each optimization problem is quite small relative to an image-then-detect approach. For all examples presented in Section 5, the Gauss-Newton method typically converged in well under 10 iterations. Finally, while the hypothesis test does require the solution of many small optimization problems, each  $\ell_m(\mathbf{y})$  may be evaluated independently so that complexity can in principle be kept low through a parallel implementation of the test.

## 4 The Anomaly Characterization Algorithm

Here we present a multiscale algorithm for the detection and characterization of an unknown number of anomalous areas in region  $A$  based on a sequence of MAP hypothesis tests. Two classes of tests are employed in this algorithm where each test takes as input the current estimate of the anomaly structures in the form of a geometry matrix,  $\hat{\mathbf{B}}$ , and an estimated amplitude vector  $\hat{\mathbf{a}}$ . The overall algorithm is a sequence of *localization hypothesis tests* which “zoom in” on the true anomalous regions coupled with *pruning tests* designed to eliminate false alarms thereby controlling the amount of work required to identify true anomalous structures.

In Fig. 7, a pseudo-code listing of the algorithm is displayed. The algorithm begins with a single localization test in which anomalous activity is identified in a collection of relatively large size areas in region  $A$ . After this coarse-scale detection procedure, the basic processing step is comprised of a single localization-type test followed by a sequence of one or more pruning tests.

The algorithm terminates when successive null hypotheses are chosen for both a localization test and then a pruning test; that is, when it is determined that no further subdivision is warranted and no structures should be removed from the current configuration. Essentially, this approach grows a tree of hypothesized anomaly configurations where the root of the tree is a single anomaly over all of region  $A$  and the leaves are either new anomaly structures or increasingly localized estimates of current anomalies.

A key feature of this procedure is the specification of the non-negative penalty terms,  $\pi_j$ , in (9) (or equivalently the prior probabilities  $p_j = e^{-\pi_j}$ ) to control the manner in which the tree is both grown (i.e. new anomalies are added or old ones further refined) as well as pruned (i.e. structures are removed from processing.) For example, hypotheses may be penalized to a greater or lesser degree depending on features such as the number of anomalies associated with  $\mathbf{B}_m$ , the spatial distribution of these structures, or geometric characteristics such as area or shape. Thus, in their ability to mathematically capture prior expectations and constraints, the  $\pi_j$  play the same role in the hypothesis testing framework as that of a smoothness regularizer in a more traditional imaging-type problem [10]. In the anomaly detection context, the property of smoothness is not of concern, but rather geometric characteristics of the anomaly distribution are the natural features for which prior models are required. The remainder of this section is devoted to a description of both classes of hypothesis tests.

#### 4.1 The Localization Hypothesis Test

Given an estimated configuration of anomalies in region  $A$ , the spatial refinement hypothesis test used here indicates either that no further localization is deemed necessary or further refinement is warranted in which case only a single structure in the current anomaly collection is further subdivided. Formally, suppose that  $N$  regions in  $A$  have been identified as containing anomalies so that  $\hat{\mathbf{B}}$  contains  $N$  columns,  $\hat{\mathbf{b}}_j$ ,  $j = 1, 2, \dots, N$ . Localization to smaller regions of  $A$  is obtained

through the use of a test with  $8N + 1$  hypotheses. The first hypothesis corresponds to retaining the current  $\widehat{\mathbf{B}}$  indicating that no further localization is warranted. Under the next 8 hypotheses, the first column of  $\widehat{\mathbf{B}}$  is replaced by one or more columns corresponding to anomalies with support over a subset of the region corresponding to  $\widehat{\mathbf{b}}_1$ . This same construction is repeated with  $\widehat{\mathbf{b}}_1$  replaced by  $\widehat{\mathbf{b}}_j$  for  $j = 2, 3, \dots, N$ . Thus, the non-null hypotheses correspond to  $8N$  ways in which anomalies can be better localized by perturbing only one of the  $N$  elements of  $\widehat{\mathbf{B}}$ .

In Fig. 4, the decompositions of  $\widehat{\mathbf{b}}_j$  of interest in this paper are displayed. The first six choices correspond to the true anomaly being located in the top, bottom, left, right, vertical middle, and horizontal middle of the region given by  $\widehat{\mathbf{b}}_j$ . The last two decompositions allow for the presence of two anomalies in  $\widehat{\mathbf{b}}_j$ : one in the top and one in the bottom or one in the left and one in the right. By choosing one of these last two hypotheses, we are able to identify multiple structures in  $A$ .

As discussed in Section 3, we associate with each hypothesis a matrix,  $\mathbf{B}_m$  for  $m = 0, 1, \dots, 8N$  where  $\mathbf{B}_0 = \widehat{\mathbf{B}}$  and for  $m > 0$ ,  $\mathbf{B}_m$  is equal to  $\widehat{\mathbf{B}}$  with one column replaced either by a single column, as is the case for the first six hypotheses in Fig. 4, or two column vectors for the last two hypotheses. In either case, we define  $\widetilde{\mathbf{b}}_m$  to be the column(s) in  $\mathbf{B}_m$  different from those in  $\widehat{\mathbf{B}}$ .

In addition to specifying  $\mathbf{B}_m$ , the MAP test requires that the penalty terms  $\pi_m$ , also be specified. For the cases where  $\widetilde{\mathbf{b}}_m$  contains one column, we begin by defining a collection of quantities  $\tilde{\pi}_m$  as

$$\tilde{\pi}_0 = 10^\sigma f_0 \left( \max_k |\widehat{\mathbf{b}}_k| \right) \quad (11)$$

and for  $m > 0$

$$\tilde{\pi}_m = \begin{cases} +\infty & |\widetilde{\mathbf{b}}_m| < 4 \text{ or } \epsilon(\widetilde{\mathbf{b}}_m) > 4 \\ 10^\sigma \text{card}(\widetilde{\mathbf{b}}_m) f_1(|\widetilde{\mathbf{b}}_m|) & \text{otherwise} \end{cases} \quad (12)$$

where  $|\widetilde{\mathbf{b}}_m|$  is the number of non-zero pixels in  $\widetilde{\mathbf{b}}_m$ ;  $\epsilon(\widetilde{\mathbf{b}}_m)$  is the ratio of the length of longer to the shorter side of the anomaly and provides a measure of the eccentricity of the structure;  $\text{card}(\widetilde{\mathbf{b}}_m)$  is the number of columns in  $\widetilde{\mathbf{b}}_m$  (in this case one); and  $f_0(|\mathbf{b}_m|)$  and  $f_1(|\mathbf{b}_m|)$  are plotted in Fig. 5.

When  $\tilde{\mathbf{b}}_m$  contains two columns, the number of nonzero pixels in each column are equal and the geometric structures of the underlying anomalies are the same (only their positions in  $A$  differ), so that there is no ambiguity in assuming that  $|\tilde{\mathbf{b}}_m|$  and  $\epsilon(\tilde{\mathbf{b}}_m)$  refer to the first column in these cases.

Finally,  $\pi_m$  are obtained from  $\tilde{\pi}_m$  via

$$\pi_m = \tilde{\pi}_m - \log \left( \sum_{k=0}^{8N} \tilde{\pi}_k \right) \quad (13)$$

thereby ensuring that  $\sum_{m=0}^{8N} p_m = \sum_{m=0}^{8N} e^{-\pi_m} = 1$ .

The parameter  $\sigma$  in (11) and (12) determines the overall magnitude of the penalty terms in the hypothesis test. Thus, selection of  $\sigma$  is analogous to the determination of the regularization parameter in many imaging type inverse problems [5, 7]. For  $\sigma$  too small the prior information will have negligible impact on the hypothesis test while an excessively large value of  $\sigma$  negates the influence of the data. In this paper,  $\sigma$  is chosen using the following somewhat *ad hoc* approach

$$\sigma = \left\lfloor \log_{10} \frac{1}{2} \|\mathbf{y} - \mathbf{h}(\mathbf{1}\hat{\mathbf{a}})\|_{\mathbf{R}^{-1}}^2 \right\rfloor - 1 \quad (14)$$

where  $\mathbf{1}$  is the indicator function over all of region  $A$ ,  $\hat{\mathbf{a}}$  is the ML amplitude for an anomaly with this structure, and  $\lfloor x \rfloor$  is the greatest integer less than or equal to  $x$ . Thus,  $\sigma$  is selected so that the largest of the  $\tilde{\pi}_m$  is roughly an order of magnitude less than the size of the covariance weighted error between the data and the initial estimated anomaly structure used in the algorithm.

Fig. 5 and eqs. (11) and (12) demonstrate that the penalty terms are chosen to reflect prior expectations regarding the sizes, shapes and number of anomalies in region  $A$ . The structure of  $f_0(|\mathbf{b}_m|)$  indicates that as the size of the largest structure in  $\hat{\mathbf{B}}$  increases, the penalty for *not* performing a decomposition increases thereby encouraging the decomposition of larger structures. The first condition in (12) is used to ensure that no structures are identified if they contain few than four pixels or if they are too long and fat or tall and thin. That is, we require anomalies to have a minimum area and not be excessively elongated. The second condition in (12) and the form of  $f_1(|\mathbf{b}_m|)$  indicate that the penalty increases as the number of columns in  $\tilde{\mathbf{b}}_m$  increases and the size

of the structures decreases. Thus, the prior expectations serve to prevent the identification of many small anomalies in region  $A$ . Clearly, depending on the application, other prior modeling choices are both possible and interesting. The selection made here is useful for the inverse conductivity problem under consideration and provides a successful demonstration of the utility of prior modeling in an hypothesis testing framework for capturing information not commonly accessible to traditional regularization schemes.

## 4.2 The Pruning-Type Hypothesis Test

The second class of hypothesis tests used here are designed to eliminate from consideration previously identified regions which are unlikely to actually contain anomalies. Given an estimated geometry matrix  $\hat{\mathbf{B}}$  with  $N$  columns, the anomaly removal hypothesis test consists of  $N + 1$  hypotheses with  $\mathbf{B}_0 = \hat{\mathbf{B}}$  and for  $m = 1, 2, \dots, N$ ,  $\mathbf{B}_m$  equals  $\hat{\mathbf{B}}$  with the  $m$ th column removed. Thus, at most one structure is eliminated per hypothesis test.

With  $\hat{\mathbf{b}}_m$  the column of  $\hat{\mathbf{B}}$  removed for the  $m$ th test,  $\tilde{\pi}_m$  for the pruning-type hypothesis test is

$$\tilde{\pi}_m = \begin{cases} 10^\sigma f_2(N) & m = 0 \\ 10^\sigma f_3(|\hat{\mathbf{b}}_m|) & m = 1, 2, \dots, N \end{cases} \quad (15)$$

where  $f_2(N)$  and  $f_3(|\hat{\mathbf{b}}_m|)$  are shown in Figs. 5(a) and (b),  $\sigma$  is chosen as in (14), and  $\pi_m$  are obtained from  $\tilde{\pi}_m$  using (13). As seen in Fig. 5, the form of  $f_2(N)$  indicates an increasing penalty as the number of estimated structures rises. Also, Fig. 5(a) shows that  $\tilde{\pi}_m$  increases as the size of the anomaly being removed from consideration rises thereby penalizing against the premature removal of large-scale structures.

## 4.3 Initial Localization

For the algorithm presented in this section to detect and localize relatively small anomalies, during the initial processing stages the correct large scale structures which overlap the true, smaller sized anomalies must be chosen for further processing. In [11], this coarse-scale localization issue

was analyzed in depth for a decision-theoretic anomaly detection algorithm applied to a scattering problem identical to the one considered here but using a linearized model rather than the non-linear model of Section 2. There, it was shown that the probability of a coarse scale miss was unacceptably high. In that case, performance was improved simply by retaining a greater number of coarse-scale regions for further decomposition than would otherwise be the case.

With these results in mind, in this paper we address the problem of coarse scale localization by modifying the hypothesis set used in the first localization test. Specifically, we supplement the decompositions of region  $A$  presented in Fig. 4 with those shown in Fig. 6 making for a 15-ary initial hypothesis test rather than the 8-ary test described in Section 4.1. Additionally, after performing this expanded localization-type test, we retain for further processing *all other* areas in region  $A$  not part of the chosen decomposition. For example, if the selection labeled “Horizontal middle two” is chosen from Fig. 6, then *four* anomalies are identified for further processing: the two square structures in the middle of region  $A$  as well as the two long, narrow regions at the top and bottom. While alternate means may be developed for addressing this issue of coarse scale localization, the method described here balances the conflicting requirements of an exhaustive, fine scale search over region  $A$  for likely targets and the desire to limit the computational overhead of the algorithm.

## 5 Examples

We consider the problem of characterizing conductivity anomalies given data from the scattering experiments described in Table 1. Region  $A$  is  $100\text{ m} \times 100\text{ m}$  and discretized into a  $16 \times 16$  array of pixels. The background conductivity is  $0.1\text{ S/m}$ . The source frequencies are  $100\text{ Hz}$  (used in a cross-well configuration) as well as  $1$  and  $10\text{ kHz}$  (used to obtain information near the left edge of  $A$ ). Both receiver arrays consist of 32 equally spaced elements extending from  $z = -0.05\text{ m}$  to  $z = 100.05\text{ m}$ . The left array is located at  $x = -0.05\text{ m}$  and the right at  $x = 100.05\text{ m}$ . The six

sources are located along the line  $x = -0.05$  m equally spaced from  $z = 0.05$  m to  $z = 99.95$  m.

The anomalies in all experiments are taken to have amplitudes of 1 S/m representing a 10:1 contrast against the background conductivity. Under conditions such as these, it is well known that the Born and Rytov linearizations of the scattering physics fail to hold thereby necessitating the use of a more complete, nonlinear scattering model. While the detection algorithm is based on the extended Born approximation (EBA) described in Section 2.1, the data vectors for all simulations are generated using the exact scattering model in (3a) and (3b).

The performance of the anomaly detection algorithm is assessed using a collection of different anomaly configurations under a variety of signal to noise conditions. For the model  $\mathbf{y}_i = \mathbf{h}_i(\mathbf{g}) + \mathbf{n}_i$  with  $\mathbf{y}_i \in \mathbb{R}^{N_i}$  and  $\mathbf{n}_i \sim \mathcal{N}(0, r_i^2 \mathbf{I})$ , the signal-to-noise ratio (SNR) in decibels is defined as

$$SNR_i = 10 \log_{10} \left\{ \frac{[\mathbf{h}_i(\mathbf{g})]^T [\mathbf{h}_i(\mathbf{g})]}{N_i r_i^2} \right\}. \quad (16)$$

For each anomaly configuration performance is examined at SNRs of 40, 20, 10, and 0 dB. Given an anomaly and SNR combination, the detection algorithm is executed for 10 independent realizations of the measurement noise process with the results of these runs forming the basis for our evaluation.

The information for each anomaly configuration is analyzed both quantitatively as well as qualitatively. From a quantitative perspective, we are concerned with the detection and false alarm probabilities associated with the algorithm. An anomaly is classified as having been detected if the  $\hat{\mathbf{B}}$  matrix produced at the last stage of the algorithm contains at least one structure which overlaps the true anomaly structure. All structures in the final  $\hat{\mathbf{B}}$  matrix not corresponding to detections are taken to false alarms. The statistic of interest here is the rate of false alarms per unit area defined as number of falsely identified pixels divided the total number of pixel in  $A$ . Finally, the results of the anomaly detection algorithm are compared against those of an imaging-type inverse scattering algorithm described in [13]. This reconstruction procedure requires the solution of a nonlinear least squares problem. The method of [13] is based on a form of the Levenburg-Marquardt algorithm

and makes use of the EBA to lower the algorithm's computational complexity.

While it is not the objective of this paper to provide an exhaustive analysis of the computational complexity of the anomaly detection algorithm, this approach is generally more efficient than the imaging technique of [13]. All algorithms in this paper have been implemented in Matlab and executed on a dual processor, SparcStation 20 with 64Mb of internal RAM. The time required for a single run of the anomaly detection algorithm varied from 4 to about 7 hours depending on the underlying anomaly configuration and the noise conditions. For the imaging algorithm, 50 iterations of the Levenburg-Marquardt technique were used to process each data set at which point the algorithm had converged to a minimum of the least-squares cost function. In contrast to the detection-based approach, the imaging algorithm required between 24 and 36 hours of processing.

### 5.1 Experiment I: A Single Anomaly

The first anomaly configuration of interest here is shown in Fig. 7(a) and consists of a single structure located near the middle of region *A*. As discussed in [11–13], for inverse problems of the type considered in this paper, the ability to resolve anomalous behavior is greatest near either vertical edge of region *A* where the sources and receivers are located. For locations in the middle of the region, it becomes increasingly difficult to detect structures or resolve fine scale detail in the conductivity profile. Finally, the structure in Fig. 7(a) is aligned on the underlying grid of  $16 \times 16$  pixels in such a manner that it is not possible for a decomposition scheme based upon the hypotheses in Figs. 4 and 6 to produce an estimated anomaly which is *exactly* of the same shape as the truth. Thus, this example represents a particularly challenging anomaly configuration for the algorithm.

In Fig. 7(b), the detection and false alarm probabilities are displayed for this case where the error bars indicate plus and minus two standard deviations from the sample mean. For the 40, 20, and 10 dB SNR cases, the anomaly was detected 100% of the time. At 0 dB, the detection



rate drops to 80%; however for the two cases in which the target was missed, the output of the algorithm did include an estimated anomaly located directly above the true structure. Thus, the target essentially was localized for all runs at all SNRs. Finally, the false alarm rate is at most 5% or 12 falsely identified pixels out of the 256. Sample estimates of anomaly configuration at 40dB and 10 dB SNR are presented in Fig. 8(a)–(b). The dashed lines overlaying these images indicate the region of  $A$  occupied by the true structure. We note that at both SNRs, the anomaly is detected. At 40 dB, there is also a small amplitude, falsely identified structure while the 10 dB example is free of such false alarms.

In Fig. 9, we display intermediate results of the algorithm for one trial at 0 dB. In Fig. 9(a), the estimated anomaly structure is shown after the first location test. Here we see that the algorithm has chosen to examine the 4 quarters of region  $A$ . Even at this coarse processing stage the structure with the largest amplitude is the quarter in which the true anomaly resides. In Figs. 9(b)–(e), the identified anomaly configuration is shown after one quarter, one half, three quarters and all of the processing for this trial. It is evident that well before the algorithm has terminated, the true anomaly is rather well localized. While there do exist falsely identified structures in the left side of Fig. 9(e) their amplitudes are far smaller than that of correctly identified anomaly. Thus, post-processing the results of the anomaly detection algorithm to remove structures whose magnitudes are in some sense small would further lower the false alarm rate. Finally, in Fig. 9(f) the results of the imaging algorithm are displayed for the same data set used to generate Figs. 9(a)–(e). In this case, the imaging algorithm does indicate the existence of a structure in the correct location. However, the geometric and amplitude information in this image is inferior to that generated by the detection algorithm, and as noted previously, the generation of this image took significantly longer than that of Fig. 9(e).

## 5.2 Experiment II: Two anomalies close to the left edge

In Fig. 7(c), the second anomaly configuration of interest in this paper is presented. Here we are interested in recovering the composition of two structures both of which are located close to the left vertical edge of region  $A$ . As discussed in Section 5.1, previous work has demonstrated that anomalous behavior is most easily detected near the sources and receivers. Thus, the structures in Fig. 7(c) are meant to test the ability of the detection algorithm to successfully isolate multiple anomalies.

For all trials and at all SNRs, Fig. 7(c) indicates that both structures are detected with a false alarm rate again less than 5%. Sample estimates are displayed for trials at 40 dB and 10 dB SNR respectively in Fig. 8(c)–(d). The reconstruction at 40 dB is basically identical to the true structures both in terms of geometry as well as amplitude. For the 10 dB case, both structures are correctly identified with the geometry of the lower anomaly rotated relative to that of the true structure. Again, the amplitudes are close to the true values of 1 S/m. In Fig. 10(a)–(e), the results of the detection algorithm at various stages of processing are displayed for one trial at 0 dB SNR. After the opening localization test, the most likely region for further decomposition corresponds to a subset of  $A$  containing both structures. As the algorithm proceeds, this area is further decomposed to better localize the two individual structures. Like the previous example, the falsely identified regions at the end of the algorithm are also those with the smallest amplitudes. Finally, in Fig. 10(f), the results of the imaging algorithm are displayed. For this particular data set, the imaging approach provides a severely degraded representation of the bottom structure and entirely misses the top anomaly.

## 5.3 Experiment III: Two separated anomalies

The anomaly configuration for the last example considered in this paper is shown in Fig. 7(e). From the previous experiments, the anomaly located close to the left edge of  $A$  should be relatively

easy to detect while the second structure positioned closer to the right vertical edge should be less detectable. This intuition is born out by the curves in Fig. 7(f) which show that both structures were detected for all trials at SNRs of 10, 20 and 40 dB. At 0 dB SNR, the smaller structure was identified in all 10 trials which the larger one was detected in 8 out of the 10 experiments. As in the first example, for the two cases in which this structures was not detected there was an estimated anomaly whose support was directly adjacent to that of the true structure. Finally, the false alarm rate for this example is again below 5%.

Sample results of the algorithm at 40 and 10 dB SNR are displayed in Fig. 8(e)–(f) where it is seen that at both noise levels, the structures are well localized. The amplitude of of the smaller structure is quite close to 1 while that of the second structure is about 1.5 for the 40 dB case and 2 at 10 dB. In Fig. 11, the results of one experiment at 0 dB SNR are plotted. Unlike the previous two cases, at the opening stages of the algorithm, the anomaly with the largest amplitude is located in a region not containing one of the true structures. Nonetheless, in the final estimate, the smaller structure is exactly identified and one of the estimated anomalies fully overlaps the second, larger structure. In contrast to the results of the detection algorithm, the imaging results for this problem, shown in Fig. 11(f) demonstrate that only a severely blurred representation of the left anomaly is available with essentially no information present regarding the structure on the right of region  $A$ .

## 6 Conclusion

In this paper, we have developed and tested an algorithm for the detection, localization, and characterization of a collection of anomalous areas in an overall region of interest from observations of scattered radiation obtained along the region’s boundary. Our approach employs a full, non-linear measurement model relating the anomalous structures to the observed fields and thus provides a

method for solving the target detection problem in cases where more common models based on the Born or Rytov linearizations are not valid. The algorithm makes no assumption concerning the number of anomalies present in the region or the amplitudes of these structures. While the method presented in this paper is adapted to the detection anomalies of which are rectangular in shape, as discussed in Section 2.2 the rectangular assumption is not required.

The anomaly detection algorithm is structured as a sequential hypothesis testing procedure which begins by localizing anomalous behavior to relatively large areas in the overall region. Subsequent tests are used to spatially refine the coarse-scale regions and prune away structures which are deemed unlikely to contain anomalies. Finally, we have developed a decision-theoretic scheme for the easy incorporation of prior information directly relevant to the anomaly detection problem into the spatial subdivision process. A MAP hypothesis testing framework was used to capture expectations regarding issues such as the number, shapes, and sizes of structures in the region.

The algorithm has been tested on a low frequency inverse electrical conductivity problem arising in geophysical exploration. The results in Section 5 demonstrate that high detection and low false alarm rates are obtainable even in highly noisy environments where more conventional imaging-type reconstruction algorithms fail to produce useful results. The decision-theoretic detection algorithm considered in this paper was far more computationally efficient than the full scale imaging algorithm and has a structure making it well suited for implementation in a parallel or distributed processing environment where these computational gains would be significantly enhanced.

## References

- [1] Richard A. Albanese, Richard L. Medina, and John W. Penn. Mathematics, medicine and microwaves. *Inverse Problems*, 10:995–1007, 1994.
- [2] Yoram Bresler, Jeffrey A. Fessler, and Albert Macovski. A Bayesian approach to reconstruction from incomplete projections of a multiple object 3D domain. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11(8):840–858, November 1989.

- [3] A. J. Devaney and G. A. Tsihrintzis. Maximum likelihood estimation of object location in diffraction tomography. *IEEE Trans. ASSP*, 39(3):672–682, March 1991.
- [4] P. E. Gill, W. Murry, and M. H. Wright. *Practical Optimization*. Academic, San Diego, Calif., 1981.
- [5] Gene H. Golub, Michael Heath, and Grace Wahba. Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics*, 21(2):215–223, May 1979.
- [6] Tarek M. Habashy, Ross W. Groom, and Brian R. Spies. Beyond the Born and Rytov approximations: A nonlinear approach to electromagnetic scattering. *J. Geophys. Res.*, 98(B2):1759–1775, February 1993.
- [7] Per Christian Hansen. Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Review*, 34(4):561–580, December 1992.
- [8] R. F. Harrington. *Field Computations by Moment Methods*. Macmillan, New York, 1968.
- [9] David Isaacson. Distinguishability of conductivities by electrical current computed tomography. *IEEE Trans. on Medical Imaging*, MI-5(2):91–95, June 1986.
- [10] Mark R. Luetttgen, W. Clem Karl, and Alan S. Willsky. Efficient multiscale regularization with applications to the computation of optical flow. *IEEE Trans. Image Proc.*, 3(1):41–64, 1994.
- [11] Eric L. Miller and Alan S. Willsky. Multiscale, statistical anomaly detection analysis and algorithms for linearized inverse scattering problems.
- [12] Eric L. Miller and Alan S. Willsky. Multiscale, statistically-based inversion scheme for the linearized inverse scattering problem. *IEEE Trans. Geosc. Remote Sens.*, 34(2):346–357, March 1996.
- [13] Eric L. Miller and Alan S. Willsky. Wavelet-based methods for the nonlinear inverse scattering problem using the Extended Born Approximation. *Radio Science*, 31(1):51–67, Jan–Feb 1996.
- [14] Leon Peters Jr., Jeffrey J. Daniels, and Jonathan D. Young. Ground penetrating radar as a subsurface environmental sensing tool. *Proc. IEEE*, 82(12):1802–1822, December 1994.
- [15] David J. Rossi and Alan S. Willsky. Reconstruction from projections based on detection and estimation of objects—parts I and II: Performance analysis and robustness analysis. *IEEE Trans. on ASSP*, ASSP-32(4):886–906, August 1984.
- [16] Ken Sauer and Bede Liu. Image reconstruction from a limited number of projections using multiple object detection/estimation. In *ICASSP 90*, volume 4, pages 1861–1864, 1990.
- [17] Alon Schatzberg, Anthony J. Devaney, and Alan J. Witten. Estimating target location from scattered field data. *Signal Processing*, 40:227–237, 1994.
- [18] Carlos Torres-Verdín and Tarek M. Habashy. Rapid 2.5-D forward modeling and inversion via a new nonlinear scattering approximation. *Radio Sci.*, 29(4):1051–1079, July–August 1994.
- [19] Harry L. Van Trees. *Detection, Estimation and Modulation Theory: Part I*. John Wiley, New York, 1968.

- [20] Song Wang, Bede Liu, and Sanjeev Kulkarni. Image reconstruction from a limited number of projections: Detection/estimation of multiple discs with unknown radii. In *Proceedings ICIP-94*, volume II, pages 854–858. IEEE, IEEE Computer Society Press, 1994.

## 7 Figure and Table Captions

**Figure 1** Inverse conductivity problem of interest in this paper.

**Figure 2** General structure of anomalous regions of interest in this paper. The magnitudes,  $a_1$  and  $a_2$  of the two anomalies shown here are proportional to the color of the corresponding rectangles.

**Figure 3** Pseudo-code for anomaly detection algorithm

**Figure 4** Hypothesized regions of support used in localization-type hypothesis tests.

**Figure 5** Penalty functions used in construction of prior models for anomaly detection algorithm.

**Figure 6** Additional regions used in initial localization-type hypothesis test.

**Figure 7** Anomaly structure to be reconstructed and associated performance curves for all examples in this paper.

**Figure 8** Sample reconstructions for all anomaly configurations at 40 and 10 dB SNR. Dotted lines indicate true positions of anomalies

**Figure 9** Reconstructions for anomaly in Fig. 7(a) in (a)–(e) and estimate obtained using imaging algorithm of [13]. Dotted lines indicate true positions of anomalies

**Figure 10** Reconstructions for anomaly in Fig. 7(c) in (a)–(e) and estimate obtained using imaging algorithm of [13]. Dotted lines indicate true positions of anomalies.

**Figure 11** Reconstructions for anomaly in Fig. 7(e) in (a)–(e) and estimate obtained using imaging algorithm of [13]. Dotted lines indicate true positions of anomalies.

**Table 1** Data set definitions for observation processes of interest in the paper.

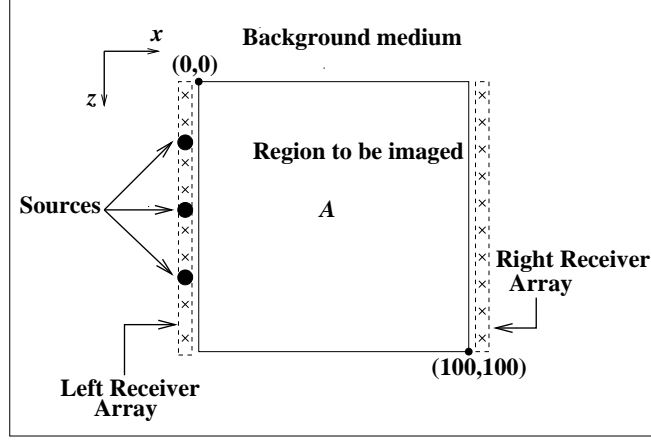


Figure 1:

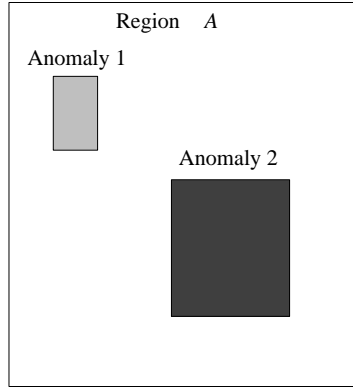


Figure 2:

| Experiment number | Source Position | Frequency of source (kHz) | Receiver Array |
|-------------------|-----------------|---------------------------|----------------|
| 1-6               | 1-6             | $f_{LO} = 0.1$            | Right          |
| 7-12              | 1-6             | $f_{MID} = 1.0$           | Left           |
| 13-18             | 1-6             | $f_{HI} = 10.0$           | Left           |

Table 1:

```

INITIALIZE ANOMALY ESTIMATES
 $\mathbf{B}_{new} := \mathbf{1}$ ;  $\mathbf{B}_{old} := \mathbf{B}_{new}$ ;
 $\mathbf{a}_{new} := 0$ ;  $\mathbf{a}_{old} := 0$ ;

PERFORM INITIAL LOCALIZATION
 $[\mathbf{B}_{new}, \mathbf{a}_{new}] := extended\_localization\_test(\mathbf{B}_{old}, \mathbf{a}_{old})$ ;
 $\mathbf{B}_{old} := \mathbf{B}_{new}$ ;  $\mathbf{a}_{old} := \mathbf{a}_{new}$ ;

START PROCESSING LOOP
done_all := false;
while (done_all  $\neq$  true) do
    PERFORM A LOCALIZATION TEST.
     $[\mathbf{B}_{new}, \mathbf{a}_{new}] := localization\_test(\mathbf{B}_{old}, \mathbf{a}_{old})$ ;

    IF NO MORE LOCALIZATION, THEN SET FLAG.
    OTHERWISE UPDATE  $B$  AND CONTINUE.
    if (identical_anomalies( $\mathbf{B}_{new}, \mathbf{B}_{old}$ ))
        no_localize := true;
    else
         $\mathbf{B}_{old} := \mathbf{B}_{new}$ ;  $\mathbf{a}_{old} := \mathbf{a}_{new}$ ;
        no_localize := false;
    endif

    PRUNE AWAY FALSE ALARMS ONE STRUCTURE AT A TIME
    done_prune := false;
    while (done_prune  $\neq$  true) do
         $[\mathbf{B}_{new}, \mathbf{a}_{new}] := prune\_test(\mathbf{B}_{old}, \mathbf{a}_{old})$ ;
        if (identical_anomalies( $\mathbf{B}_{new}, \mathbf{B}_{old}$ )  $\vee$  (number_anomalies( $\mathbf{B}_{new}$ ) == 0))
            done_prune := true;
            no_prune := true;
        else
             $\mathbf{B}_{old} := \mathbf{B}_{new}$ ;  $\mathbf{a}_{old} := \mathbf{a}_{new}$ ;
        endif
    end

    IF NO NEW LOCALIZATION AND NO PRUNING THEN TERMINATE
    if ((no_localize == true)  $\wedge$  (done_prune == true))
        all_done == true;
    endif
end

```

Figure 3:



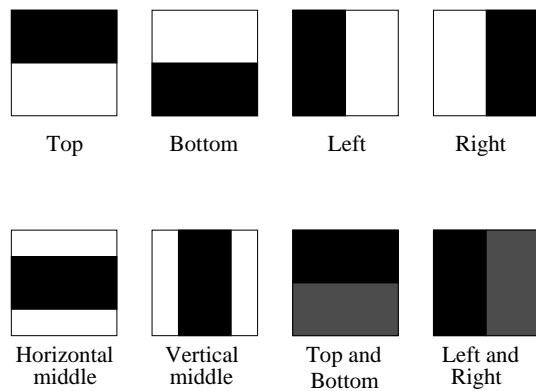


Figure 4:

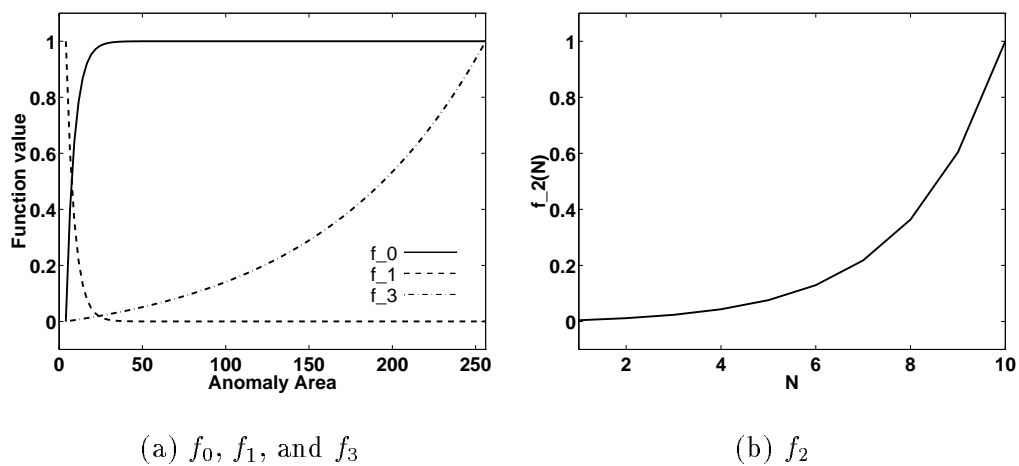


Figure 5:

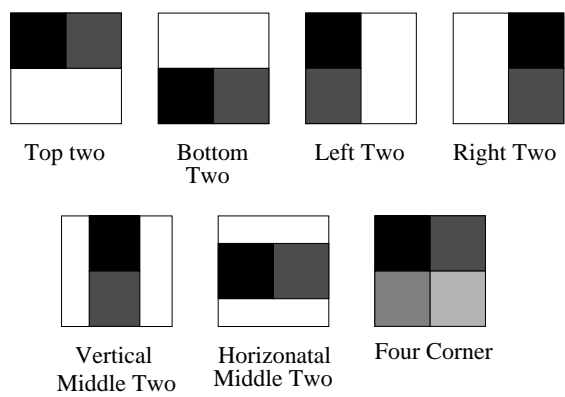
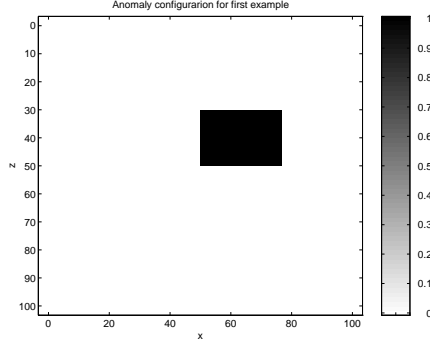
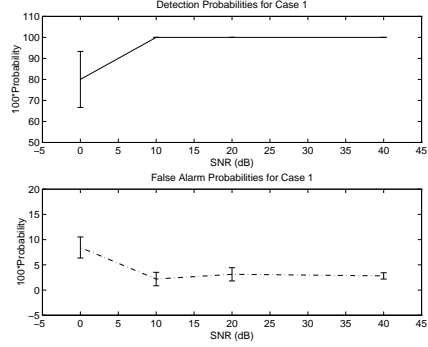


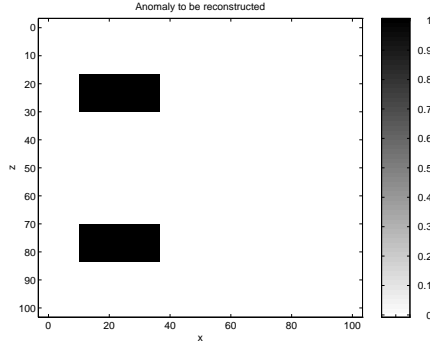
Figure 6:



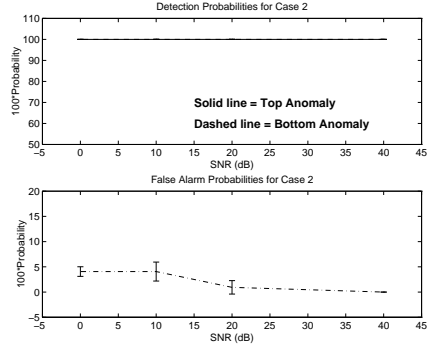
(a) Anomaly distribution for first example.



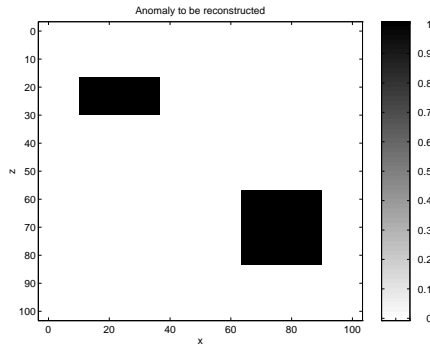
(b) Detection (top) and false alarm (bottom) probabilities for first example



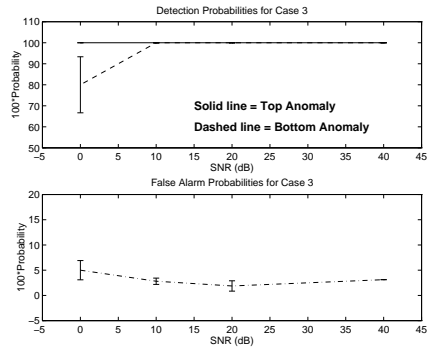
(c) Anomaly distribution for second example.



(d) Detection (top) and false alarm (bottom) probabilities for second example.

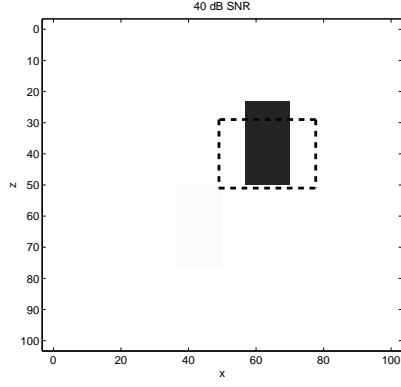


(e) Anomaly distribution for third example.

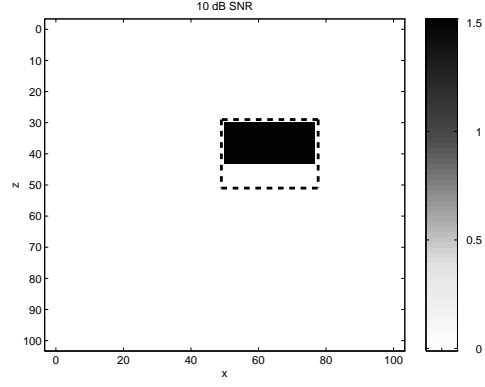


(f) Detection (top) and false alarm (bottom) probabilities for third example

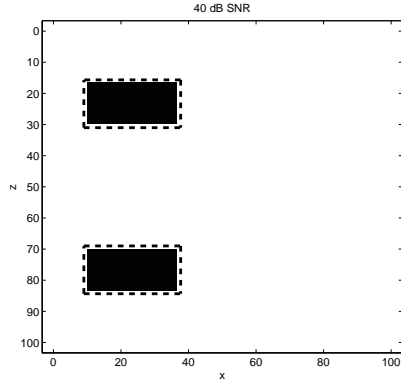
Figure 7:



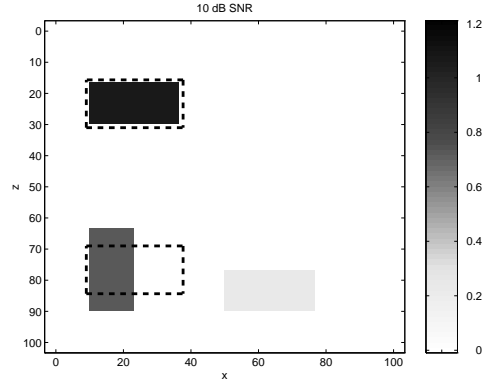
(a) Sample reconstruction at 40 dB for anomaly in Fig. 7(a)



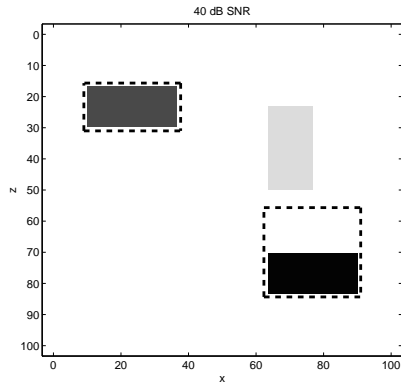
(b) Sample reconstruction at 10 dB for anomaly in Fig. 7(a)



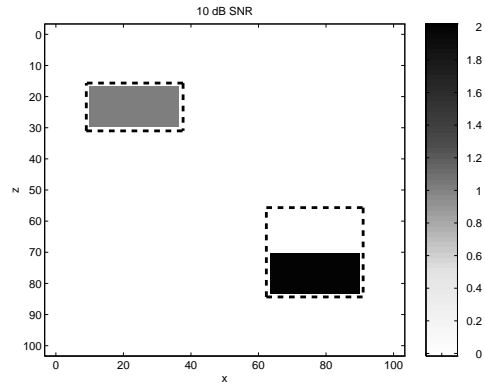
(c) Sample reconstruction at 40 dB for anomaly in Fig. 7(c)



(d) Sample reconstruction at 10 dB for anomaly in Fig. 7(c)

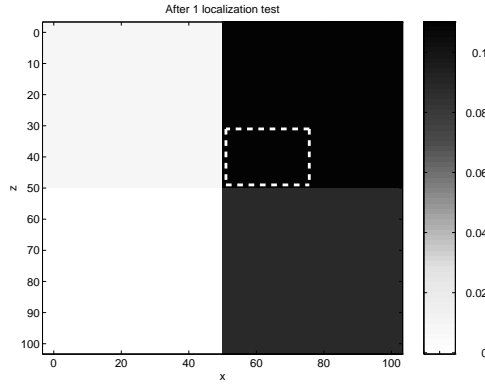


(e) Sample reconstruction at 40 dB for anomaly in Fig. 7(e)

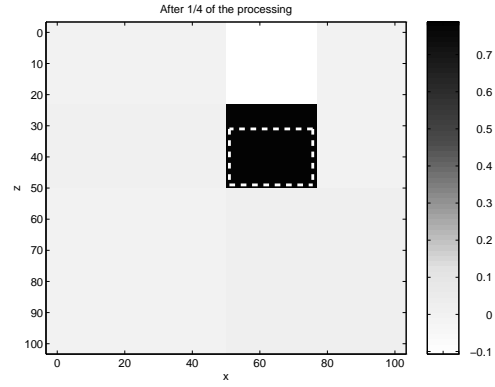


(f) Sample reconstruction at 10 dB for anomaly in Fig. 7(e)

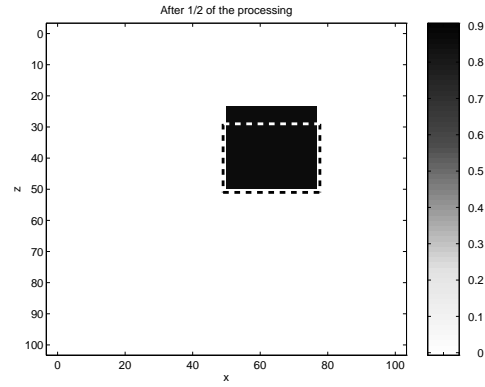
Figure 8:



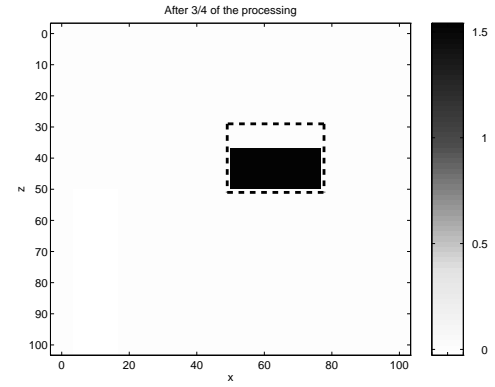
(a) Estimate after initial localization test for anomaly in Fig. 7(a)



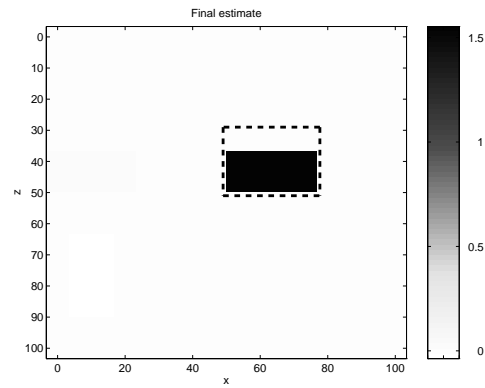
(b) Estimate after 1/4 of detection algorithm for anomaly in Fig. 7(a)



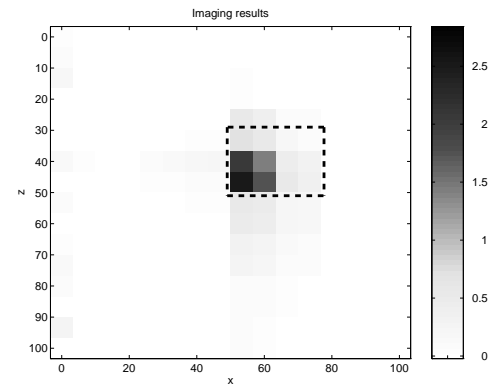
(c) Estimate after 1/2 of detection algorithm for anomaly in Fig. 7(a)



(d) Estimate after 3/4 of detection algorithm for anomaly in Fig. 7(a)

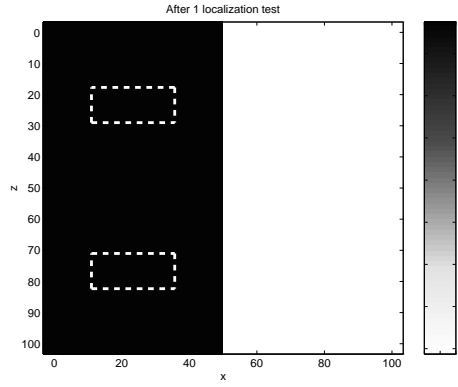


(e) Final estimate for anomaly in Fig. 7(a)

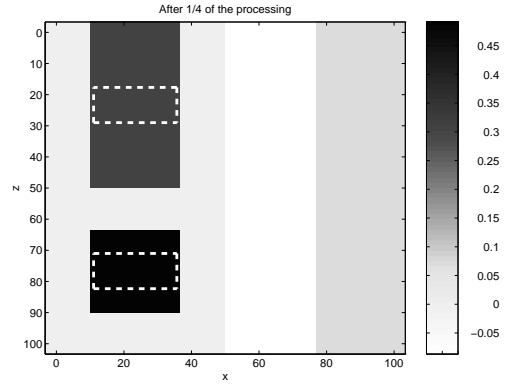


(f) Estimate after initial localization test for anomaly in Fig. 7(a)

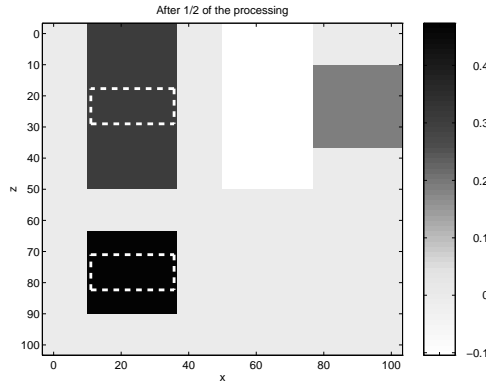
Figure 9:



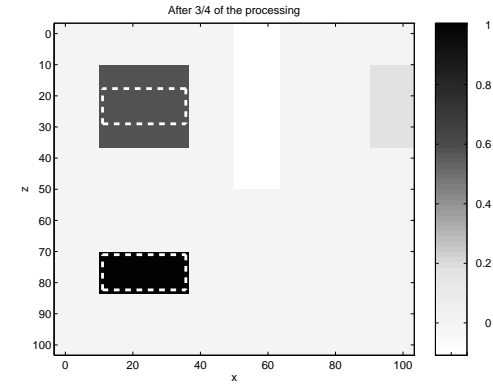
(a) Estimate after initial localization test for anomaly in Fig. 7(c)



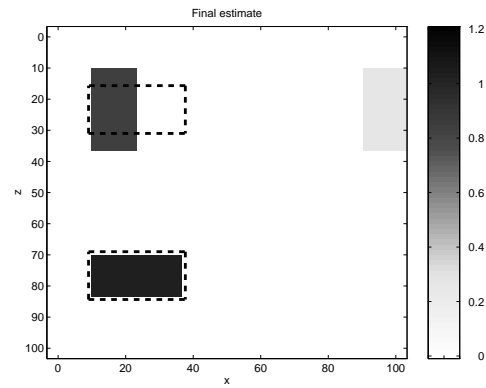
(b) Estimate after 1/4 of detection algorithm for anomaly in Fig. 7(c)



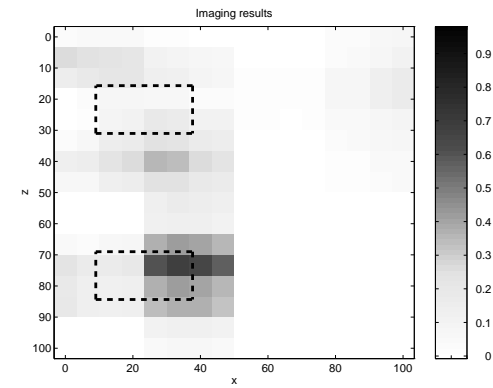
(c) Estimate after 1/2 of detection algorithm for anomaly in Fig. 7(c)



(d) Estimate after 3/4 of detection algorithm for anomaly in Fig. 7(c)

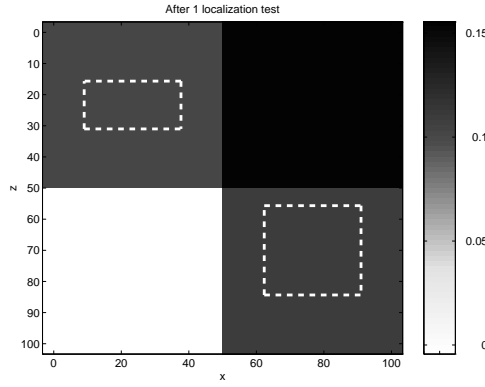


(e) Final estimate for anomaly in Fig. 7(c)

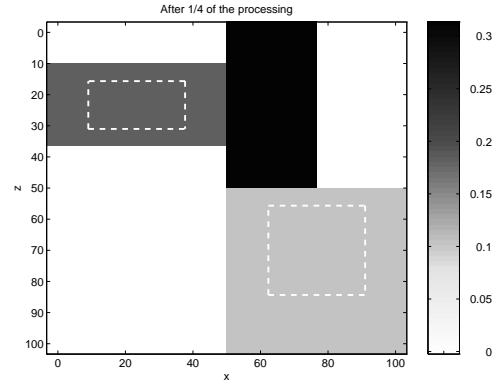


(f) Estimate after initial localization test for anomaly in Fig. 7(c)

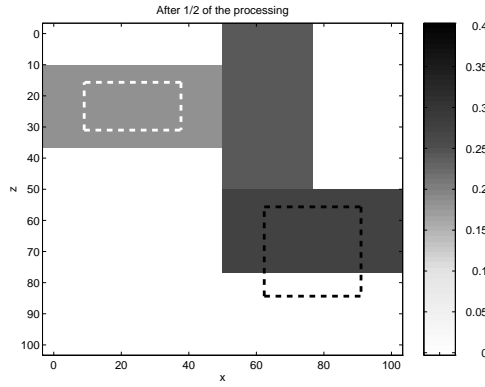
Figure 10:



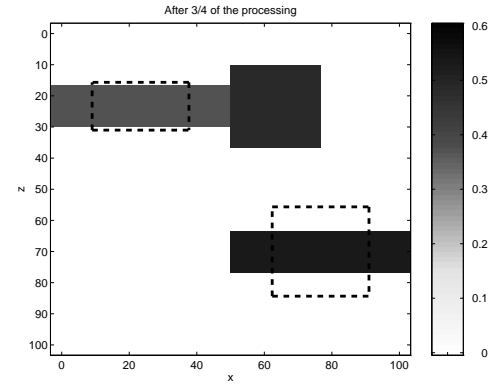
(a) Estimate after initial localization test for anomaly in Fig. 7(e)



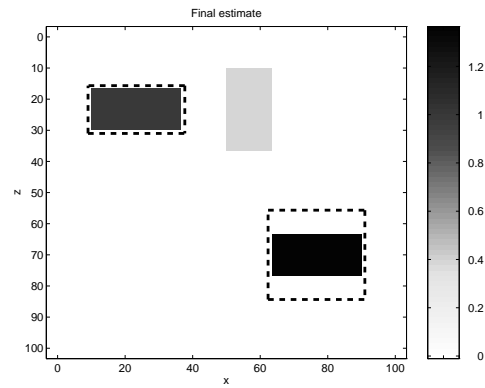
(b) Estimate after 1/4 of detection algorithm for anomaly in Fig. 7(e)



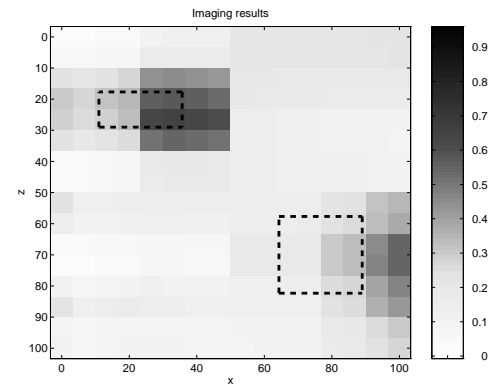
(c) Estimate after 1/2 of detection algorithm for anomaly in Fig. 7(e)



(d) Estimate after 3/4 of detection algorithm for anomaly in Fig. 7(e)



(e) Final estimate for anomaly in Fig. 7(e)



(f) Estimate after initial localization test for anomaly in Fig. 7(e)

Figure 11: