Bayesian Approach With the Maximum Entropy Principle in Image Reconstruction From Microwave Scattered Field Data

Maï Khuong Nguyen and Ali Mohammad-Djafari

***Abstract*—Microwave imaging is of great interest in medical applications owing to its high sensitivity with respect to dielectric properties. It allows detection of very small inhomogeneities. The image reconstruction employing the microwave inverse scattering consists of reconstructing the image of an object from the scattered field measured behind the object. This reconstruction runs up against the nonuniqueness of the solution of the inverse scattering problem. In this communication, we propose to solve the ill-posed inverse problem by a statistical regularization method based on the Bayesian maximum *a posteriori* (MAP) estimation where the principle of maximum entropy (ME) is used for assigning the *a priori* laws. The results obtained demonstrate the power and potential of this method in image reconstruction.**

I. Introduction

In an imaging system using inverse scattering, the object is illuminated by an electromagnetic wave of known characteristics. A multidimensional inverse scattering of wavefields is used as a means of obtaining information on the geometric structure or the interior parameters of the scatterer in a homogeneous environment. By definition, the scattered field is the difference between the fields in absence and in presence of the object. This field behaves as if it were created by an equivalent current source $\vec{J}$ which depends on the difference of the dielectric properties (conductivity and permittivity) between the object and the background medium [1], of which the amplitude J is:

$$\begin{cases} J \neq 0 & \text{inside the object} \\ J = 0 & \text{outside the object.} \end{cases}$$

Therefore, the object is entirely represented by the equivalent current distribution.

The aim of our study is to determine the shape and the location of the object, hence to determine the equivalent current distribution. This distribution is related to the scattered field $\vec{E}$ by the following expression:

$$\vec{E}(\vec{r}) = -i\omega_0\mu \int\int\int_{(v)} \overline{\overline{G}}(\vec{r},\vec{r'})\vec{J}(\vec{r'})dv + \vec{B} \qquad (1)$$

Manuscript received May 8, 1992; revised January 9, 1993 and December 25, 1993. The associate editor responsible for coordinating the review of this paper and recommending its publication was Y. Censor.

A. Mohammad-Djafari is with the Laboratoire des Signaux et Systèmes (CNRS-ESE-UPS), École Supérieure d'Électricité, Plateau de Moulon, 91192 Gif-sur-Yvette Cédex, France.

M. K. Nguyen is with the Université de Cergy-Pontoise, 95014 Cergy-Pontoise Cédex, France.

IEEE Log Number 9401070.

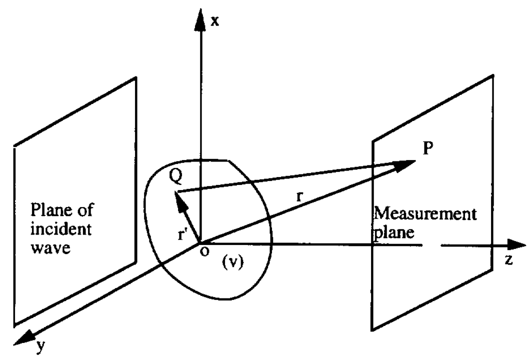

Fig. 1. Principal schema of the imaging system.

where

$\vec{E}$ is a vectorial and complex representation of the scattered field,

$\vec{J}$ is the vector which represents the equivalent current distribution,

$\overline{\overline{G}}(\vec{r},\vec{r'})$ is a Green's dyadic which is given by

$$\overline{\overline{G}}(\vec{r},\vec{r'}) = \left[\overline{\overline{I}} - \frac{1}{k^2}\overline{\overline{\text{grad}}}_r\overrightarrow{\text{grad}}_{r'}\right]\Phi$$

Φ is a Green's scalar function having the following form

$$\Phi = \frac{e^{-ik|\vec{r}-\vec{r'}|}}{4\pi \mid \vec{r}-\vec{r'} \mid}$$

$\overline{\overline{I}}$ is the identity dyadic,

k is the wave number,

$\vec{r}(x,y,z)$ is the vector which marks an observation point,

$\vec{r'}(x',y',z')$ is the vector which locates a point inside the source,

v is the volume of the object,

ω_0 is the angular frequency of the time harmonic incident wave,

μ is the magnetic permeability of the medium, and

$\vec{B}$ is a vectorial and complex representation of the measurement noise.

The principle of the imaging system can be seen in Fig. 1.

The reconstruction of the equivalent current distribution from the scattered field is in general an ill-posed inverse problem because of nonuniqueness in the solution. This nonuniqueness exactly reflects the ambiguity of a physical phenomenon in the problems of "Inverse Source" and of "Inverse Scattering": indeed, there is an infinity of current densities which can create the same field, or there are several objects which can produce the same scattered field. Therefore, the scattered field data are not sufficient to specify a unique reconstruction [2], [9], [10]. Furthermore, in reality the measurement data are often corrupted by noise. These factors lead us to search a solution which is not only unique but also stable. Analogous problems have been discussed in [6], [15], [20], [23], [24], [30].

Integral equation (1) when discretized becomes

$$\mathbf{e} = G\mathbf{j} + \mathbf{b} \tag{2}$$

where

$\mathbf{e} = \{e_1, e_2, \ldots, e_M\}$ is the data vector whose element e_k is the scattered field sample at the kth point on the measurement surface, and M is the number of measurement points;

$\mathbf{j} = \{j_1, j_2, \ldots, j_N\}$ is the unknown vector whose element j_k represents the unknown equivalent current distribution at the kth pixel of the object, and N is the number of pixels;

$\mathbf{G} = \{g_{m,n}, m = 1, \ldots, M, n = 1, \ldots, N\}$ is a matrix of size $M \times N$ whose elements are calculated from the Green's dyadic $\overline{\overline{G}}(\vec{r}, \vec{r'})$;

$\mathbf{b} = \{b_1, b_2, \ldots, b_M\}$ is the noise vector whose elements represent the noise at the different points of measurement.

The objective is to obtain a unique and stable estimate $\hat{\mathbf{j}}$ of the object $\mathbf{j}$ from the data e.

However, the ill-posed nature of integral equation (1) leads to the matrix G of (2) being in general ill-conditioned or even singular. Thus, a naive solution $\hat{\mathbf{j}} = G^{-1}\mathbf{e}$ is unrealistic.

The generalized-inverse solution or the least-squares solution with minimum norm:

$$\widehat{\mathbf{j}} = \arg\min_{\mathbf{j}} \{||\mathbf{e} - G\mathbf{j}||^2\} = (G^t G)^{-1} G^t \mathbf{e}$$

seems to be a reasonable choice (this solution is unbiased and of minimum variance from a statistical standpoint), but it is usually unacceptable because it resolves only the uniqueness but not the stability of the solution with respect to erroneous or noisy data [8].

Since the ill-posed characteristics of the inverse problem are due to the lack of information, one needs to complete the data with all other available information. The existing regularization methods consist of combining *a priori* knowledge with the information from the data in order to yield a unique and stable solution. In such methods, a compound criterion is introduced and optimized. The *a priori* information can be in a deterministic form (positivity, etc.) or in a stochastic form (means, variances, expectations, etc.). Different *a priori* information can lead to markedly different results.

For example, if the *a priori* knowledge concerns the smoothness of the solution $\mathbf{j}$, the Tikhonov regularization can be consistently used [27], [31]. This traditional regularization method consists of defining the solution which satisfies a mixed quadratic criterion:

$$\widehat{\mathbf{j}} = \arg\min_{\mathbf{j}} \{||\mathbf{e} - G\mathbf{j}||^2 + \lambda||\mathbf{j}||^2\}$$

where λ is the regularization parameter. This solution is unique, stable, linear with respect to the data and can be calculated explicitly, but there is no guarantee that it is positive, a point which is required in many imaging applications.

A priori knowledge of the positivity of the solution $\mathbf{j}$ historically leads to the so-called "ME methods". The use of entropy is rigorously founded and applied in different situations, see [4], [5], [7], [11]-[14], [17], [18], [25], [32] among others. While the arguments for ME appear similar in several applications, the ME principle is used in different ways and the ME solutions are not mathematically equivalent. We discuss here three techniques applying the ME principle.

The first technique consists of considering the unknown as a probability distribution and of choosing, among the possible solutions of the equation $\mathbf{e} = G\mathbf{j}$, the one which maximizes the entropy

$$S = -\sum_{k=1}^{N} j_k \log j_k.$$

This problem can be solved when using Lagrange multipliers and gives a guaranteed positive and smooth solution, if it exists. Methods based on this concept are known as "classical ME" methods [21].

The second technique applying the ME principle is to consider the unknown $\mathbf{j}$ as the mean values of a random vector $\mathbf{X}$ with a probability density function (PDF) $p(x)$, i.e., $\mathbf{j} = E\{\mathbf{X}\} = \langle\mathbf{X}\rangle$ and consider the data $\mathbf{e} = Gj = G\langle\mathbf{X}\rangle$ as the linear constraints on these mean values. Among the PDF $p(x)$ satisfying the data constraints we choose the one that maximizes the entropy

$$S = -\int p(x) \log p(x) dx.$$

Once $p(x)$ is determined, the solution is deduced as follows:

$$\widehat{\mathbf{j}} = E\{\mathbf{X}\} = \int x p(x) dx.$$

Methods based on this concept are referred to as "ME in mean" methods [21].

The above two methods use the principle of ME directly as a criterion for choosing a unique solution. However, they do not explicitly take noise into account.

We propose here a third technique in which the ME principle is used for assigning the *a priori* probability distribution needed in the Bayesian approach. If we can represent the *a priori* knowledge about the unknown $\mathbf{j}$ by a PDF $p(\mathbf{j})$ and the uncertainty on the data e by a conditional PDF $p(e/\mathbf{j})$, Bayes' theorem allows us to combine them in the *a posteriori* probability distribution

$$p(\mathbf{j}/\mathbf{e}) \propto p(\mathbf{j})p(\mathbf{e}/\mathbf{j})$$

and the solution can be determined by maximizing $p(\mathbf{j}/\mathbf{e})$ (the MAP criterion). Now before applying Bayes' rule, it is necessary to assign the *a priori* probability law $p(\mathbf{j})$ and the conditional PDF $p(\mathbf{e}/\mathbf{j})$ of the data. This is a difficult task since these probabilities are in general not available. It is to resolve this difficulty that we propose using the ME principle to translate our *a priori* knowledge about the object $\mathbf{j}$ and about the noise $\mathbf{b}$ into probabilistic terms if this *a priori* knowledge is in the form of statistical constraints (the expected values or the moments, etc). Methods based on this concept are called "Bayesian ME" methods [21]. In this paper, we propose a regularization method which is based on the Bayesian MAP estimation with the ME *a priori* laws: the Bayesian approach allows both the data and the *a priori* information to be incorporated into the estimation process, and the ME principle is used to translate, in a coherent way, the *a priori* knowledge into an *a priori* probability law.

In the following are presented (1) a brief description of the Bayesian approach, (2) the use of the ME principle in deducing the parametric *a priori* probability laws from the available knowledge, (3) the Bayesian ME solution, (4) the choice of the regularizing functions, (5) the simultaneous estimation of the unknowns and of the hyperparameters (regularization parameters or parameters of the *a priori* laws), and (6) the extension of the proposed method in the case of complex quantities.

II. Basics of the Proposed Method

A. Bayesian Approach

The problem we would like to solve is: "what are the best estimates of parameters that one can make from the imperfect data and the *a priori* information". We solve this problem using a Bayesian approach in which the basic steps are summarized as follows:

- Assign an *a priori* probability law $p(\mathbf{j})$ to the unknown object $\mathbf{j}$ according to the *a priori* knowledge of this latter.
- Assign a conditional probability law $p(\mathbf{e}/\mathbf{j})$ to the data according to the knowledge of the noise statistics.
- Bayes' theorem determines the *a posteriori* probability law $p(\mathbf{j}/\mathbf{e})$:

$$p(\mathbf{j}/\mathbf{e}) = \frac{p(\mathbf{j})p(\mathbf{e}/\mathbf{j})}{p(\mathbf{e})}. \tag{3}$$

  Clearly, (3) combines the *a priori* information in terms of the *a priori* probability $p(\mathbf{j})$ with the data information in terms of the probability of the data conditioned on the real solution $p(\mathbf{e}/\mathbf{j})$.
- When using the MAP criterion as a decision rule, the solution is obtained by maximizing this *a posteriori* distribution:

$$\widehat{\mathbf{j}} = \arg\max_{\mathbf{j}}\{p(\mathbf{j}/\mathbf{e})\}. \tag{4}$$

  Further information about the Bayesian approach can be found in [3], [16], [19], [28], [29] among others.

The resolution of (4) requires a calculation of two terms of (3): the *a priori* probability distribution $p(\mathbf{j})$ of the parameter to be estimated and the conditional probability distribution $p(\mathbf{e}/\mathbf{j})$ of the data (the denominator of (3) $p(\mathbf{e})$ is a constant and does not affect the optimization).

Assuming that measurement equation (1) and the noise statistics are known, determination of $p(\mathbf{e}/\mathbf{j})$ is usually straightforward.

The calculation of the first term $p(\mathbf{j})$ in (3), i.e., how to deduce an *a priori* probability $p(\mathbf{j})$ for the image from the *a priori* knowledge, is the main difficulty in the Bayesian approach. In general, the *a priori* information is not directly given in a probability form and does not yield a unique *a priori* law either. In the cases where the *a priori* knowledge takes the form of statistical constraints, we propose to use the ME principle for assigning the *a priori* probability law.

B. Principle of ME for Assigning the Probability Law

The idea of the ME principle is that if the information is not sufficient to determine uniquely a probability law, we choose, among the possible distributions satisfying the given constraints, the one which maximizes the Shannon entropy

$$\hat{p}(\mathbf{j}) = \arg\max_{p\in P}\left\{-\int p(\mathbf{j})\log p(\mathbf{j})d\mathbf{j}\right\} \tag{5}$$

where P is the set of probability distributions satisfying the constraints.

The justification for the use of ME lies in the fact that the ME distribution exactly reflects the state of *a priori* knowledge, in other words it does not introduce any other extraneous information. The ME description has been interpreted as the description most objective or maximally noncommittal with respect to missing information [5], [32]. Moreover, the ME principle translates in a coherent manner the *a priori* information into a probability law.

Indeed, let us consider the case where our *a priori* knowledge is in the form:

$$\begin{cases} E\{H(\mathbf{j})\} &= h \\ E\{S(\mathbf{j})\} &= s \end{cases} \tag{6}$$

where $H(\mathbf{j})$ and $S(\mathbf{j})$ are two known functions, h and s are constant.

Maximizing entropy (5) subjected to constraints (6) gives a solution in exponential form for the *a priori* probability law [26]:

$$p(\mathbf{j}) = \frac{1}{Z(\lambda,\mu)}\exp\left[-\lambda H(\mathbf{j}) - \mu S(\mathbf{j})\right] \tag{7}$$

where the partition function $Z(\lambda,\mu)$ is given by

$$Z(\lambda,\mu) = \int \exp\left[-\lambda H(\mathbf{j}) - \mu S(\mathbf{j})\right]d\mathbf{j}$$

and the Lagrange multipliers λ, μ are deduced from (h, s) by solving the system of equations

$$\begin{cases} -\frac{\partial \ln Z(\lambda,\mu)}{\partial\lambda} &= h \\ -\frac{\partial \ln Z(\lambda,\mu)}{\partial\mu} &= s. \end{cases} \tag{8}$$

Expression (7) represents a family of ME *a priori* laws that depend on two parameters (λ, μ). Such *a priori* probability laws are completely consistent with the *a priori* knowledge.

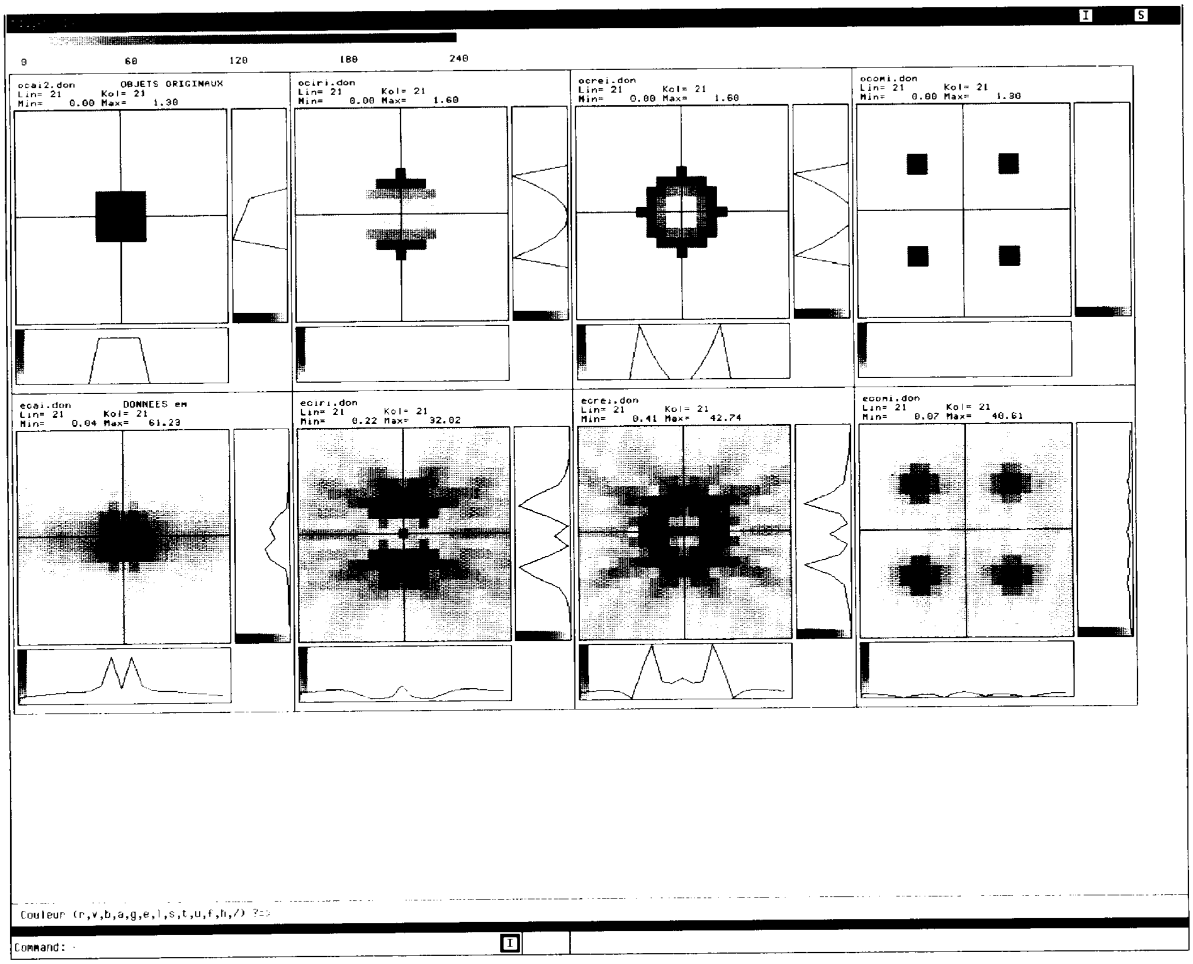

Fig. 2. The objects of different shapes {01, 02, 03, and 04} are in the first row, the modules of the corresponding complex scattered fields are in the second row.

C. Conditional Probability of the Data

Making the standard assumptions about the noise, i.e., an independent white noise with zero-mean and variance σ^2, the ME principle leads to a Gaussian distribution for the probability law of the data $p(\mathbf{e}/\mathbf{j})$:

$$p(\mathbf{e}/\mathbf{j}) \propto \exp[-Q(\mathbf{j})] \tag{9}$$

where $Q(\mathbf{j}) = [\mathbf{e}-G\mathbf{j}]^t W[\mathbf{e}-G\mathbf{j}]$ and $W = \mathrm{diag}[\frac{1}{\sigma^2}, \ldots, \frac{1}{\sigma^2}]$.

D. Bayesian ME Solution

Taking into account expressions (7) and (9) for $p(\mathbf{j})$ and $p(\mathbf{e}/\mathbf{j})$, respectively, the MAP solution (4) is equivalent to

$$\hat{\mathbf{j}} = \arg\min_{\mathbf{j}}\{F(\mathbf{j}) = Q(\mathbf{j}) + \lambda H(\mathbf{j}) + \mu S(\mathbf{j})\} \tag{10}$$

in which both noise and *a priori* information are present.

Thus, the image to be recovered is one which satisfies criterion (10). In summary, the Bayesian estimation with the ME principle gives a solution which is regularized thanks to the *a priori* knowledge introduced in the form of the regularizing functions H and S, or equivalently translated into the *a priori* probability law in a coherent way.

E. Hyperparameters

In practice, to solve problem (10), it is important to determine the hyperparameters (λ, μ). Several cases may be considered:

Firstly, (λ, μ) are given or can be directly deduced in an empirical manner from some physical conditions.

Secondly, the expectations (h, s) are known and thus (λ, μ) can be calculated from (8).

Thirdly, the most general case, where only the regularizing functions (H, S) are known but not the expectations (h, s), one must estimate (λ, μ) as well as $\hat{\mathbf{j}}$. To do this, there are two statistical methods: successive and simultaneous estimation [22]. In this paper, we propose using the simultaneous estimation because the successive estimation method requires the calculation of the marginalized likelihood which is often impracticable (the Gaussian case excepted).

F. Choice of the Functions H and S

In general, H and S can be any arbitrary function. However, if we wish the form of the *a priori* law to be independent of the scale factor of the image (i.e., it remains invariant when

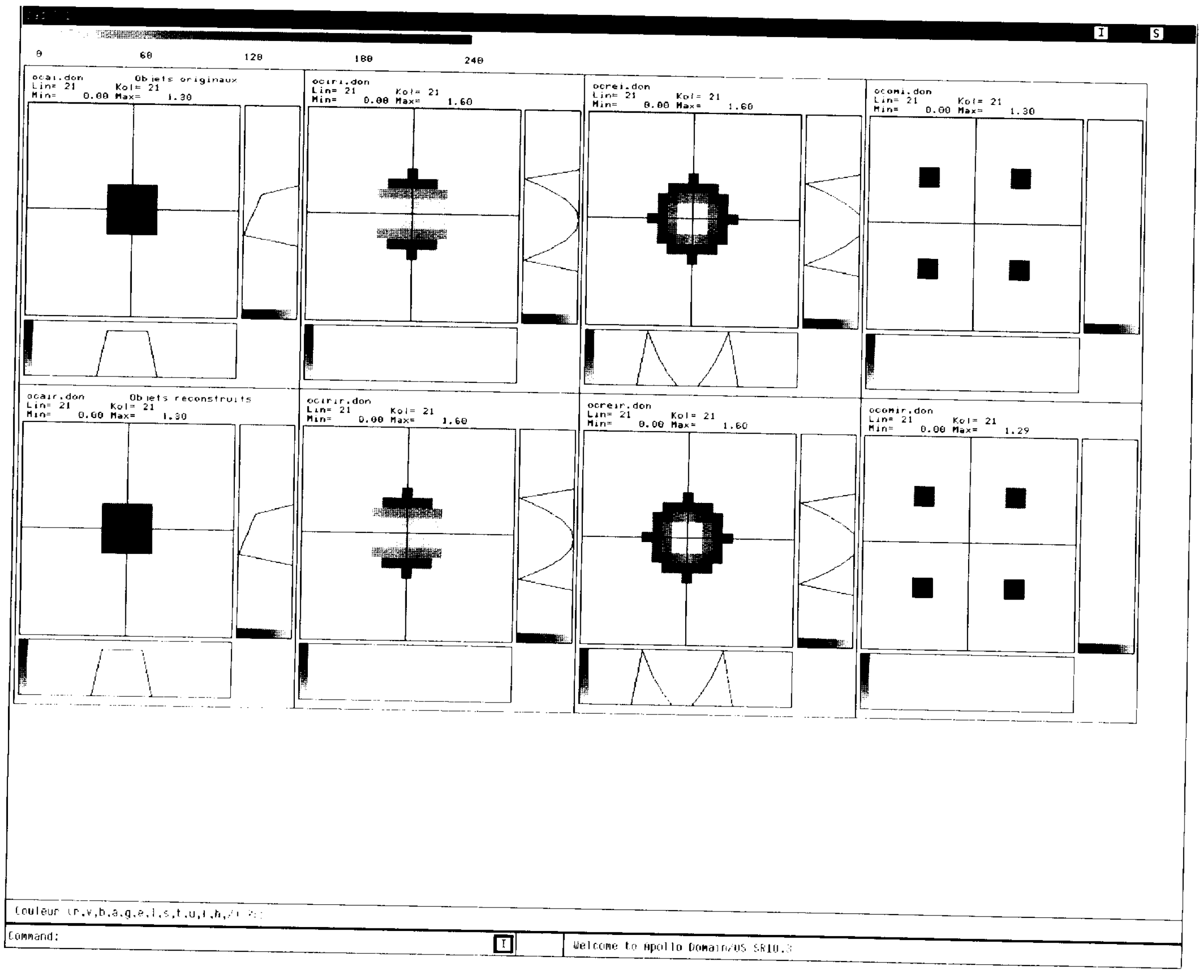

Fig. 3. Reconstruction of different inhomogeneous objects {01, 02, 03, and 04} from the simulated data without noise. The horizontal and vertical curves trace the pixel values in the corresponding coordinate axes. The different gray levels represent the range of absolute values.

the scale of the image is changed), and with the hypothesis of the independence of pixels, the admissible forms for H and S are restricted to simple combinations of power and logarithmic functions [26]. In our study, we have considered two typical forms of (H, S): if the pixel values must be positive

$$H(\mathbf{j}) = \sum_{k=1}^{N} \log j_k \quad \text{and} \quad S(\mathbf{j}) = \sum_{k=1}^{N} j_k \tag{11}$$

else

$$H(\mathbf{j}) = \sum_{k=1}^{N} j_k^2 \quad \text{and} \quad S(\mathbf{j}) = \sum_{k=1}^{N} j_k. \tag{12}$$

Case (11) yields the following *a priori* law $p(\mathbf{j})$:

$$p(\mathbf{j}) = \frac{1}{Z(\lambda, \mu)} \exp\left[-\lambda \sum_{k=1}^{N} \log j_k - \mu \sum_{k=1}^{N} j_k\right] \tag{13}$$

which is a multivariate Gamma probability distribution.

Case (12) leads to the Gaussian *a priori* law:

$$p(\mathbf{j}) = \frac{1}{Z(\lambda, \mu)} \exp\left[-\lambda \sum_{k=1}^{N} j_k^2 - \mu \sum_{k=1}^{N} j_k\right]. \tag{14}$$

The reason for choosing these forms of *a priori* laws is that in these cases it is possible to obtain analytical and explicit relations between (λ, μ) and the two first moments of the probability law. Indeed, denoting the mean: $m = E\{j\}$ and the variance: $v = E\{(j - m)^2\}$, we obtain the following useful relations:

$$\begin{cases} \lambda &= \frac{v - m^2}{v} \\ \mu &= \frac{m}{v} \end{cases} \tag{15}$$

for the Gamma *a priori* law, and

$$\begin{cases} \lambda &= \frac{1}{2v} \\ \mu &= -\frac{m}{v} \end{cases} \tag{16}$$

for the Gaussian *a priori* law.

We have developed an algorithm for calculating (m, v) from the estimated image $\hat{\mathbf{j}}$ at each iteration by the moment method

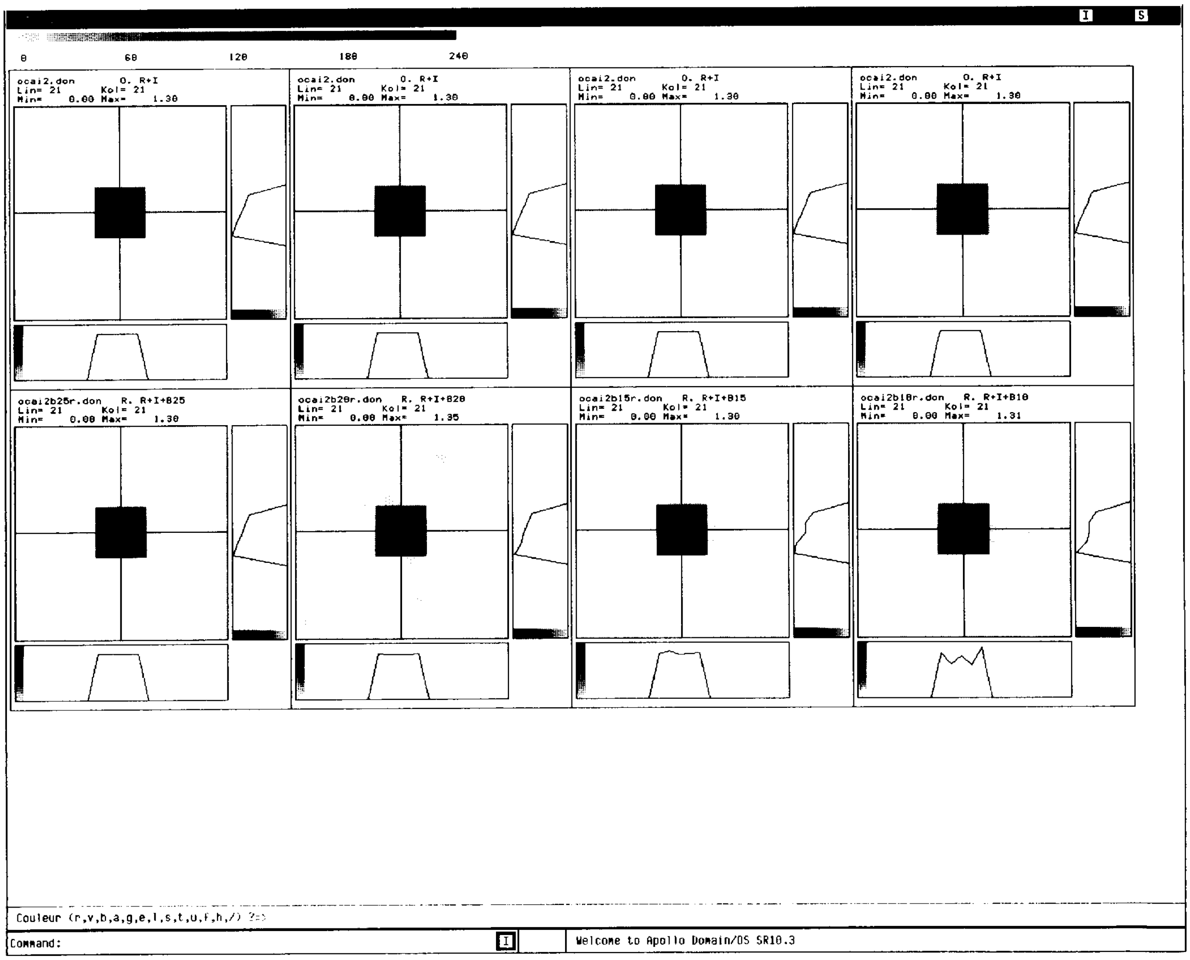

Fig. 4. Reconstruction of the inhomogeneous square object 01 from the simulated data for different levels of noise: $\frac{S}{N}$ = {25dB, 20dB, 15dB, 10dB} from left to right, respectively.

(MM). The hyperparameters (λ, μ) are thus calculated from relations (15) and (16).

G. Case of Complex Pixel Values

In general, the data $\mathbf{e}$, the noise $\mathbf{b}$ and the matrix G are complex quantities. The object $\mathbf{j}$ can be real or complex.

If $\mathbf{j}$ is real, and assuming that the real and imaginary parts of the noise are independent, (2) becomes:

$$\begin{bmatrix} \mathbf{e}^R \\ \mathbf{e}^I \end{bmatrix} = \begin{bmatrix} G^R \\ G^I \end{bmatrix} [\mathbf{j}] + \begin{bmatrix} \mathbf{b}^R \\ \mathbf{b}^I \end{bmatrix} \tag{17}$$

and the application of the method is unchanged except that the dimensions of the vectors in (13) are doubled.

If $\mathbf{j}$ is complex, the reconstruction method can be extended when the real and imaginary parts of $\mathbf{j}$ are considered as independent. Then (2) can be rewritten as

$$\begin{bmatrix} \mathbf{e}^R \\ \mathbf{e}^I \end{bmatrix} = \begin{bmatrix} G^R & -G^I \\ G^I & G^R \end{bmatrix} \begin{bmatrix} \mathbf{j}^R \\ \mathbf{j}^I \end{bmatrix} + \begin{bmatrix} \mathbf{b}^R \\ \mathbf{b}^I \end{bmatrix}. \tag{18}$$

In this case, the functions H and S in (11) and (12) are defined for real and imaginary parts of pixel values assumed positive:

$$H(\mathbf{j}) = \sum_{k=1}^{N} \left(\log j_k^R + \log j_k^I\right) \quad \text{and} \quad S(\mathbf{j}) = \sum_{k=1}^{N} \left(j_k^R + j_k^I\right) \tag{19}$$

otherwise:

$$H(\mathbf{j}) = \sum_{k=1}^{N} \left[(j_k^R)^2 + (j_k^I)^2\right] \quad \text{and} \quad S(\mathbf{j}) = \sum_{k=1}^{N} \left(j_k^R + j_k^I\right). \tag{20}$$

The calculation is performed on both the real and imaginary parts, the previous development of the reconstruction method remaining unchanged.

H. Summary of the Bayesian ME Method

The proposed method leads to theoretical solution (10) which is obtained as a result of the Bayesian MAP estimation with the ME *a priori* law: the Bayesian approach is used to combine both the information from the data and the *a priori*

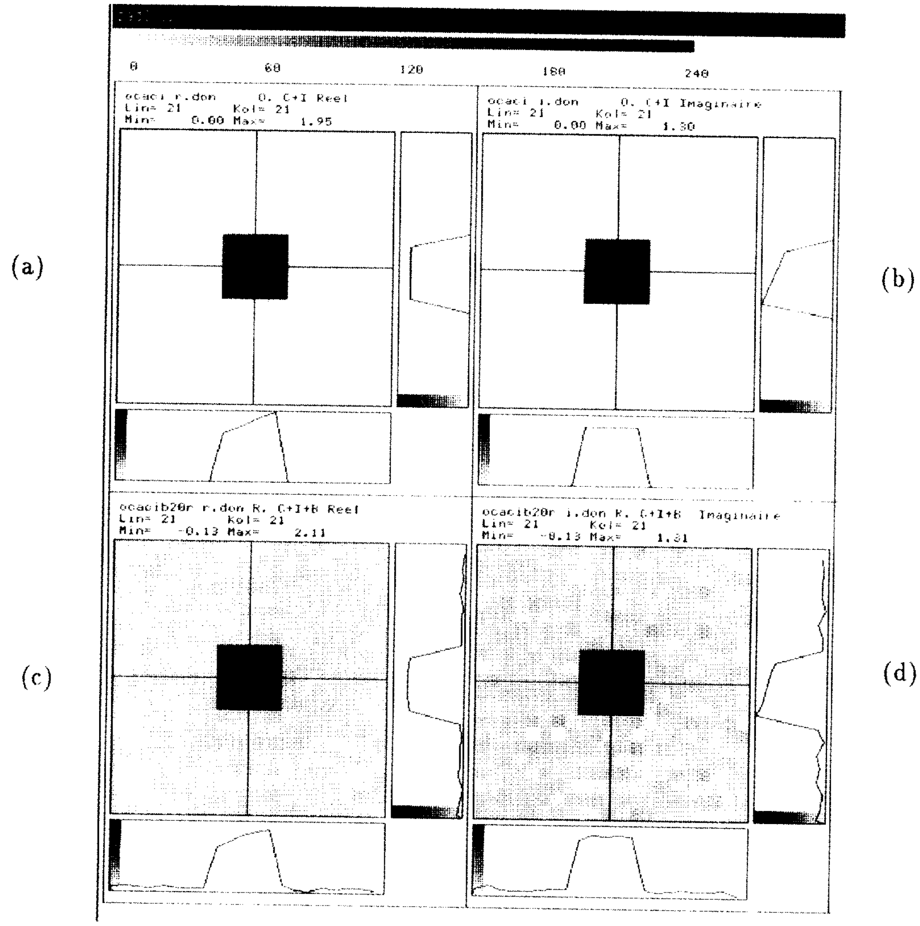

Fig. 5. Reconstruction of a complex object from the simulated data for the signal-to-noise $\frac{S}{N}$ = 20dB. The first row represents the real part (a) and imaginary part (b) of the original object. The second row shows the real part (c) and the imaginary part (d) of the reconstructed object.

knowledge in the estimation process. In this approach the decision rule is the MAP and the ME principle is used in the first step for deducing the *a priori* probability from the *a priori* information.

The practical resolution of (10) is performed by a simultaneous estimation of the solution $\widehat{\mathbf{j}}$ and the hyperparameters (λ, μ). We have developed an algorithm in which (λ, μ) are calculated from the moments (m, v) at each iteration, and the unknown $\widehat{\mathbf{j}}$ is estimated according to criterion (10) by the conjugate gradient method.

III. Numerical Simulation Process

According to the above mentioned method, the reconstruction simulation process is carried out as follows: from the definition of an object $(\mathbf{j})$, the scattered field $(\mathbf{e})$ is calculated and the reconstruction $(\widehat{\mathbf{j}})$ of the object is performed according criterion (10).

The reconstruction procedure is illustrated for a two-dimensional case. Let us use as incident wave an x-axis polarized plane-wave propagating in the z-axis, i.e., the amplitude of incident wave E_i is given by

$$E_i(\vec{r}) = E_{ix} \exp(-ikz) \text{ and } E_{iy} = E_{iz} = 0.$$

The frequency of the incident wave is 2.5 GHz.

The scattered field is calculated in a plane parallel to the plane of the incident wave and placed behind the object (Fig. 1), at a measurement distance of about 10 wavelengths. The calculation of one component of relation (1), E_x and J_x for example, is as follows:

$$E_x(\vec{r}) = -i\omega_0\mu \int\int\int_{(v)} \left(\Phi - \frac{1}{k^2}\frac{\partial^2\Phi}{\partial x \partial x'}\right) J_x(\vec{r'}) dv. \quad (21)$$

The reconstruction simulation process is detailed below:

Step 1: the object is constructed from a number of pixels traversed by an equivalent current distribution $\mathbf{j} = \{j_1, j_2, \ldots, j_N\}$ with j_k different from zero inside the object and equal to zero outside the object.

Step 2: the scattered field is calculated by (17) on a plane surface and the values are stored as vector $\mathbf{e}$. To simulate the measurement noise, Gaussian random noise is generated with a given signal-to-noise ratio and added to vector $\mathbf{e}$.

TABLE I
THE VALUES D OF RECONSTRUCTION WITHOUT NOISE ($\frac{S}{N}$ = 25 dB)

Objects	01	02	03	04
D of Fig. 3	1.142e-07	1.832e-06	1.056e-05	127e-06

TABLE II
THE VALUES D OF RECONSTRUCTION FOR DIFFERENT LEVELS OF NOISE

Levels of Noise	25 dB	20 dB	15 dB	10 dB
D of Fig. 4	1.142e-07	3.606e-05	6.418e-05	114.9e-05

TABLE III
THE VALUES D OF RECONSTRUCTION OF A COMPLEX OBJECT FOR ($\frac{S}{N}$ = 20 dB)

Levels of Noise	20 dB
D of Fig. 5	7.222e-02

Step 3: minimizing the function $\{F(\mathbf{j}) = Q(\mathbf{j}) + \lambda H(\mathbf{j}) + \mu S(\mathbf{j})\}$ provides the estimated object $\hat{\mathbf{j}}$. This step uses the conjugate gradient technique.

IV. RESULTS

In order to validate the method, the algorithm was tested with objects of different shapes such as: square ($O1$), circular ($O2$), hollow ($O3$) and compound ($O4$) objects.

A number of reconstructions were performed: homogeneous and inhomogeneous objects reconstructed from the simulated data for different levels of noise (the signal-to-noise ratio $\frac{S}{N}$ varies between 25 dB and 10 dB). The data are presented in Fig. 2, the results in Fig. 3 and 4.

Fig. 2: represents the different objects ($O1$, $O2$, $O3$ et $O4$) in the first row and the modules of the corresponding scattered fields are in the second row.

Fig. 3: shows the reconstruction of inhomogeneous different objects ($O1$, $O2$, $O3$ et $O4$) from the simulated data without noise.

Fig. 4: gives the reconstruction of the inhomogeneous square object $O1$ from the simulated data for different levels of noise:

$$\left(\frac{S}{N} = 25 \text{ dB}, \frac{S}{N} = 20 \text{ dB}, \frac{S}{N} = 15 \text{ dB}, \frac{S}{N} = 10 \text{ dB}\right).$$

Fig. 5: shows the reconstruction of a complex object from the simulated data for the signal-to-noise $\frac{S}{N}$ = 20dB. The first row represents the real and imaginary parts of the original object. The second row shows the real and imaginary parts of the reconstructed object.

In Figs. 3, 4, and 5, the horizontal and vertical curves trace the pixel values in the corresponding coordinate axes. The different grey levels represent the range of absolute values. It should be noted that there is excellent agreement between the original and the reconstructed objects.

In order to estimate the reconstruction quality, we have calculated the mean energy of reconstruction error for the image:

$$D = \frac{\sum_{k=1}^{N} (j_k - \hat{j}_k)^2}{N}$$

where N is the number of image pixels.

Tables I to III show the quality measurements of reconstruction in absence and in presence of noise:

Furthermore, the absolute original and reconstructed (maximal and minimal) values are visualized and show the similarity between them.

V. CONCLUSION

The numerical simulation results show the high quality of the Bayesian ME reconstruction. Even in the case of much noise or missing data, the absolute values of the pixels can be changed but the shape of the object is always well reconstructed. This interesting result can be explained by the fact that in the case of missing or noisy data, a good reconstruction is performed as a result of the appropriate regularization introduced in a coherent way by the ME principle.

Although there exists Fourier relations between the scattered field and the equivalent current distribution, and although most of the image reconstruction methods often use Fourier data [1], [23], [24], we have proposed here a statistical method which directly operates in the spatial domain. We can thus eliminate errors coming from the numerical calculation of direct and inverse Fourier transformations and from the truncation of signals.

Concerning the computational load, the correct solution is obtained after a reasonable number of iterations (about ten iterations for the gradient conjugate technique). In comparison with conventional Fourier methods, which need one direct Fourier transformation (FT) and one inverse FT, each iteration takes about the same CPU time as a FT. Thus the quality of the reconstruction is acquired at the expense of the computational time (about 10 times longer).

Nevertheless, the high quality of the reconstructions and the robustness of the solution under missing or noisy data demonstrate the power and potential of the proposed method for image reconstruction.

ACKNOWLEDGMENT

This work has been performed in the LSS-Inverse Problem Group under the guidance of Professor Guy Demoment.

REFERENCES

[1] M. Baribaud, "Review article: microwave imagery: analytical method and maximum entropy method," *J. Phys. D: Appl. Phys. 23*, pp. 269–288, 1990.
[2] W. M. Boerner and C. Y. Chan, "Inverse method in electromagnetic imaging," *Medical Applicat. of Microwave Imaging*, L. E. Larsen and J. H. Jacobi, eds., M.S.E.E. 1985.
[3] T. M. Cannon, H. J. Trussell, and B. R. Hunt, "Comparison of image restoration methods," *Appl. Optics*, vol. 17, no. 21, pp. 3384–3390, Nov. 1978.
[4] Y. Censor, "Finite series-expansion reconstruction methods," *Proc. IEEE*, vol. 71, pp. 409–419, Mar. 1983.
[5] Y. Censor and A. Lent, "Optimization of "log x" entropy over linear equality constraints," *SIAM J. Cont. and Optimization*, vol. 25, no. 4, pp. 921–932, July 1987.
[6] W. C. Chew and Y. W. Wang, "Reconstruction of two-dimensional permittivity distribution using the distorted Born iterative method," *IEEE Trans. Medical Imaging*, vol. 9, pp. 218–225, June 1990.
[7] G. J. Daniell and S. F. Gull, "Maximum entropy algorithm applied to image enhancement," *IEE Proc.*, vol. E-127, pp. 170–172, 1980.

[8] G. Demoment, "Reconstruction and restoration: overview of common estimation structures and problems," *IEEE Trans. Acoust., Speech and Signal Processing*, vol. 37, pp. 2024–2035, Dec. 1989.

[9] A. J. Devaney and G. C. Sherman, "Nonuniqueness in inverse source and scattering problems," *IEEE*, EP-30, no. 5, 1982.

[10] M. Fischer and K. J. Langenberg, "Limitations and defects of certain inverse scattering theories," *IEEE*, AP-32, no. 10, 1984.

[11] B. R. Frieden, "Restoring with maximum likelihood and maximum entropy," *J. Opt. Soc. Am.*, vol. 62, no. 4, pp. 511–518, Apr. 1972.

[12] B. R. Frieden, "Statistical models for the image restoration problem," *Comput. Graph. Image Proces.*, vol. 12, pp. 40–59, 1980.

[13] B. R. Frieden and D. C. Wells, "Restoring with maximum entropy. III. Poisson sources and backgrounds," *J. Opt. Soc. Am.*, vol. 68, no. 1, pp. 93–103, Jan. 1978.

[14] S. F. Gull and J. Skilling, "Maximum entropy method in image processing," *IEE Proc.*, F-131, pp. 646–659, 1984.

[15] P. Hua, E. J. Woo, J. G. Webster, and W. J. Tompkins, "Iterative reconstruction method using regularization and optimal current pattern in electrical impedance tomography," *IEEE Trans. Medical Imaging*, vol. 10, pp. 621–628, Dec. 1991.

[16] B. R. Hunt, "Bayesian methods in nonlinear digital image restoration," *IEEE Trans. Comput.*, vol. C-26, pp. 219–229, Mar. 1977.

[17] E. T. Jaynes, "Prior probabilities," *IEEE Trans. Syst. Sci. and Cybern.*, vol. SSC-4, pp. 227–241, Sept. 1968.

[18] E. T. Jaynes, "On the rationale of maximum entropy methods," *IEEE Proc.*, vol. 70, no. 9, pp. 939–952, 1982.

[19] T. Lei and W. Sewchand, "Statistical approach to X-ray CT imaging and its applications in image analysis—Part I: Statistical analysis of X-ray CT imaging,—Part II: A new stochastic model-based image segmentation technique for X-ray CT image," *IEEE Trans. Medical Imaging*, vol. 11, pp. 53–69, Mar. 1992.

[20] M. Moghaddam and W. C. Chew, "Nonlinear two-dimensional velocity profile inversion using time domain data," *IEEE Trans. Geosci. and Remote Sensing*, vol. 30, pp. 147–156, Jan. 1992.

[21] A. Mohammad-Djafari, "Maximum entropy and linear inverse problems: a short review," *Maximum Entropy and Bayesian Methods*, A. Mohammad-Djafari and G. Demoment, Eds., New York: Kluwer Academic Publishers, 1993.

[22] A. Mohammad-Djafari, "On the estimation of hyperparameters in Bayesian approach of solving inverse problems," in *Proc. IEEE ICASSP-93*, vol. V, pp. 495–498, 1993.

[23] A. Mohammad-Djafari and G. Demoment, "Maximum entropy Fourier synthesis with application to diffraction tomography," *Appl. Optics*, vol. 26, no. 10, pp. 1745–1754, 1987.

[24] A. Mohammad-Djafari and G. Demoment, "Maximum entropy reconstruction in X-ray and diffraction tomography," *IEEE Trans. Medical Imaging*, vol. 7, no. 4 pp. 345–354, 1988.

[25] A. Mohammad-Djafari and G. Demoment, "Utilisation de l'entropie dans les problèmes de restauration et de reconstruction d'images," *Traitement du Signal*, vol. 5, no. 4, pp. 235–248, 1988.

[26] A. Mohammad-Djafari and J. Idier, "Maximum entropy prior laws of images and estimation of their parameters," , *Maximum Entropy and Bayesian Methods*, W. T. Grandy, Jr. and L. H. Schick, Eds., pp. 285–293, New York: Kluwer Academic Publishers, 1991.

[27] A. Tikhonov and V. Arsenin, *Solutions of Ill-Posed Problems*. Washington, DC: Winston, 1977.

[28] H. J. Trussell, "Notes on linear image restoration by maximizing the *a posteriori* probability," *IEEE Trans. Acoust., Speech and Signal Processing*, vol. ASSP-26, pp. 174–176, Apr. 1978.

[29] "The relationship between image restoration by the maximum *a posteriori* method and a maximum entropy method," *IEEE Trans. Acoust., Speech and Signal Processing*, vol. ASSP-28, pp. 114–117, Feb. 1980.

[30] G. A. Tsihrintzis and A. J. Devaney, "Maximum likelihood estimation of object location in diffraction tomography, Part II: Strongly scattering objects," *IEEE Trans. Signal Processing*, vol. 39, pp. 1466–1470, June 1991.

[31] S. Twomey, "On the numerical solution of Fredholm integral equation of the first kind by the inversion of the linear system produced by quadrature," *J. Ass. Comput.*, vol. 10, no. 6, pp. 97–101, March 1962.

[32] S. J. Wernecke and L. R. D'addario, "Maximum entropy image reconstruction," *IEEE Trans. Comput.*, vol. C-26, pp. 351–364, Apr. 1977.