Microwave imaging: characterization of unknown dielectric or conductive materials

Olivier Féron*, Bernard Duchêne* and Ali Mohammad-Djafari*

*Laboratoire des Signaux et Systèmes, CNRS–SUPÉLEC–UPS
Supélec, Plateau de Moulon, 91192 Gif-sur-Yvette, France

Abstract. Microwave imaging problem consists in reconstructing unknown objects from measurements of the scattered field that result from their interaction with a known interrogating wave. This problem is nonlinear and ill-posed. The main classical methods to solve this inverse problem are based on the linearization of the model (by using Born or Rytov approximation) or work directly on the nonlinear mapping. In both cases the inverse problem is solved by minimizing a cost functional that can be, in a Bayesian estimation framework, interpreted as a Maximum a posteriori (MAP) estimate. The classical prior information introduced is a smoothness or contour preserving constraint. In this paper we propose to introduce the information that the object is composed of a finite (known) number of materials by using hierarchical Markov Random Field modeling approach. We then propose a Bayesian inversion method and compute the Posterior Mean estimate by using appropriate Markov Chain Monte Carlo (MCMC) algorithms.

Key words: Microwave imaging, Inverse Problems, Hierarchical modeling, Gibbs sampling.

INTRODUCTION

In this paper we deal with an electromagnetic inverse obstacle scattering problem. An incident field $\phi_{0,q,v}$ is created at different frequencies $q \in \{1, \ldots, N_q\}$ and different incidence views $v \in \{1, \ldots, N_v\}$. The interactions between these incident fields and the unknown object $x$ create diffracted fields $y_{q,v}$, at each frequency and each view, which is measured on the receiver positions $s \in S$. The object $x$ is characterized by its relative permittivity $\varepsilon_r(r)$ and its conductivity $\sigma(r)$ (Siemens/m) and is given by the form

$$x(r) = \varepsilon_r(r) - 1 + i \sigma(r) = x'(r) + i x'(r)$$  \hspace{1cm} (1)
In a 2D "transverse magnetic" configuration the direct model can be written with observation and state equations that are respectively

\[ y_{q,v} = G^S_{q,v} X_q \phi_{q,v}, \quad q = 1, \ldots, N_q, \quad v = 1, \ldots, N_v \]  

(2)

\[ \phi_{q,v} = \phi_{0q,v} + G^D_q X_q \phi_{q,v}, \quad X_q = \text{diag}(x_q) \]  

(3)

where \( G^S_{q,v} \) and \( G^D_q \) are respectively \((\text{card}(S) \times \text{card}(D))\) and \((\text{card}(D) \times \text{card}(D))\) complex matrices associated to the Green’s kernels, \( \phi_{q,v} \) is the total field inside \( D \) \( (\phi_{q,v}(r) = \phi_{0q,v}(r) + y_{q,v}(r), \ r \in D) \), and \( x_q \) is defined by

\[ x_q(r) = x^f(r) + i \frac{x^i(r)}{\Omega_q} \]  

(4)

where \( \Omega_q \) is a constant depending on the frequency \( q \). The inverse problem consists in retrieving the contrast function \( x(r) \), \( r \in D \), from the measured scattered fields \( y_{q,v}(s), s \in S \), through the inversion of the two coupled equations (2) and (3). This problem is nonlinear, since in the observation equation (2) the total field \( \phi_{q,v} \) inside \( D \) shows up, and this total field also depends on the contrast \( x \) as indicated by the state equation (3). A first class of methods to solve this nonlinear problem consists in linearizing the model (by Born or Rytov approximations [6]) and then use classical optimization methods. A second class of methods works directly on the nonlinear model and often consists in jointly estimating \( x \) and \( \phi \) from the measurements \( y = \{y_{q,v}, \ q = 1, \ldots, N_q, \ v = 1, \ldots, N_v\} \). For example in the called "Modified Gradient" method [4] the estimation of \( x \) and \( \phi \) is done by minimizing a criterion that is a combination of two mean square error terms corresponding to the two coupled equations (2) and (3). By this way the authors use the "biconvexity" of the criterion due to the bilinearity of the model with respect to \( x \) and \( \phi \), in order to alternatively minimize \( x \) and \( \phi \). Even if this kind of methods works well on simply (for example binary) objects, the criterion is not convex and then only the convergence to a local minimum is ensured. Theoretically this problem can be avoided by using stochastic sampling methods. Here we propose to use the bilinearity property of this nonlinear model by implementing an MCMC algorithm. Particularly a Gibbs sampling method allows to still handle \( x \) and \( \phi \) alternatively. However we will see that the exact sampling of \( x \) is possible if we consider the new modeling of the coupled equations

\[ y_{q,v} = G^S_{q,v} w_{q,v} + \varepsilon_{q,v}, \]  

(5)

\[ w_{q,v} = X_q \phi_{0q,v} + X_q G^D_q w_{q,v} + \eta_{q,v}, \]  

(6)

with the change of variable \( w_{q,v} = X_q \phi_{q,v} \) and where \( \varepsilon_{q,v} \) and \( \eta_{q,v} \) are assumed to be complex, independent and Gaussian noises: \( \varepsilon_{q,v} \sim \mathcal{N}(0, \rho_\varepsilon I_{2N_S}) \) and \( \eta_{q,v} \sim \mathcal{N}(0, \rho_\eta I_{2N_D}) \), where \( N_S \) and \( N_D \) are respectively the number of elements of \( S \) and \( D \). The objective of this paper is then to estimate jointly \( x \) and \( w = \{w_{q,v}, q = 1, \ldots, N_q, \ v = 1, \ldots, N_v\} \) from the measurements \( y = \{y_{q,v}, q = 1, \ldots, N_q, \ v = 1, \ldots, N_v\} \). The main classical method to solve (5) and (6) is the contrast source inversion (CSI) method [7]
that consists in minimizing the criterion

\[
J_{\text{est}}(x, w) = \frac{\sum_{q,v} \|y_{q,v} - G^S_q w_{q,v}\|^2}{\sum_{q,v} \|y_{q,v}\|^2} + \frac{\sum_{q,v} \|X_q \phi_{0,q,v} - w_{q,v} + X_q G^D_q w_{q,v}\|^2}{\sum_{q,v} \|X_q \phi_{0,q,v}\|^2},
\]

alternatively on \(x\) and \(w\). In this paper we propose to compute in a Bayesian estimation framework the Posterior Mean of the a posteriori distribution \(p(x, w | y)\)

\[
p(x, w | y) \propto p(y | w) p(w | x) p(x),
\]

where the two first terms are given by the coupled equations and the probability distributions of the noises \(\epsilon_{q,v}\) and \(\eta_{q,v}\). We now have to define the prior probability \(p(x)\).

PRIOR CHOICE

Regularization methods have already been developed to solve this inverse problem and often consist in introducing prior information that the object is smooth (Markovian type regularization term [1]) or with contours (edges preserving term [5]). In this paper we propose to introduce the prior information that the object is composed of a finite (known) number of homogeneous materials. We therefore consider a hierarchical Markov modeling on the unknown \(x\) [3], based on the introduction of a hidden variable \(z(r)\). This new variable takes discrete values \(k \in \{1, \ldots, K\}\), where \(K\) is the number of materials, and then represents a classification label. Each class \(k\) is characterized by a complex mean \(m_k = m_k^r + i m_k^i\) and two variances \(\rho_k^r\) et \(\rho_k^i\) respectively for the real and imaginary parts. By this way we can define the prior conditional probability of a pixel \(x(r)\) given its label \(z(r)\) for a Gaussian Mixture (GM) model :

\[
p(x(r)|z(r) = k, m_k, \rho_k^r, \rho_k^i) = \mathcal{N} \left( \begin{bmatrix} m_k^r \\ m_k^i \end{bmatrix}, \begin{bmatrix} \rho_k^r & 0 \\ 0 & \rho_k^i \end{bmatrix} \right),
\]

where we consider \(x(r)\) as a vector \(x(r) = [x^r(r), x^i(r)]^T\) decomposing it by its real and imaginary parts. Given the classification \(z = \{z(r), r \in D\}\), all the pixels \(x(r)\) are independent and a spatial correlation is only considered on the labels by means of a Potts Markov Random field :

\[
p(z) = \frac{1}{Q(\alpha)} \exp \left\{ \alpha \sum_{r \in D} \sum_{r' \in \mathcal{N}(r)} \delta[z(r) - z(r')] \right\},
\]

with \(\delta\) is the Kronecker delta function and where \(\mathcal{N}(r)\) is the neighborhood of \(r\) (of 4 pixels), \(Q(\alpha)\) is the partition function and \(\alpha\) represents the degree of spatial dependency of \(z\). As we assume that the different materials are quite homogeneous we fix the variances \((\rho_k^r, \rho_k^i)\) to small values \((10^{-4})\). However we have to estimate the means \(m_k\) and the variances \(\rho_\epsilon\) and \(\rho_\eta\) of noises \(\epsilon_{q,v}\) and \(\eta_{q,v}\). For easily sampling these called hyper-parameters we choose conjugate priors for them, which are Gaussian \(\mathcal{N}(\mu_k^r, \sigma_k^r)\)
As we chose conjugate priors for the variances \( \mu_k^t, \sigma_k^t \) for the means \( m_k^t \) and \( m_l^t \) of each class \( k \) and Inverse Gamma \( \Gamma (\alpha_e, \beta_e) \) and \( \Gamma (\alpha_\eta, \beta_\eta) \) for the variances \( \rho_e \) and \( \rho_\eta \). Because we have not any prior idea about the value of \( \rho_e \) and \( \rho_\eta \) we choose \( \alpha_e, \beta_e, \alpha_\eta, \beta_\eta \) that give quite flat prior distributions. Note here that if we know the characteristics of some of the materials that compose the value of \( m_1 \) and \( m_2 \) we then simply have to affect particular peaky Gaussian prior distribution on the means \( m_k^t \) and \( m_l^t \) with small variances. For example in the following the label 1 corresponds to the air and we then take a centered Gaussian distribution for \( m_1 \) with small variance.

### Gibbs Sampling Algorithm

In the following we note \( m = \{ m_k, k = 1, \ldots, K \} \) regrouping the means of all the classes. The whole set of unknowns is then \( (x, w, z, m, \rho_e, \rho_\eta) \). Here we propose to estimate the posterior mean by using a Gibbs sampling algorithm that consists in decomposing the whole set of variables into subsets and alternatively sampling the corresponding conditional probability laws. Actually the change of variable explained in the introduction will allow not only to obtain a tractable probability \( p(x \mid w, z, m, \rho_\eta) \) but also to have a complex separable distribution \( p(w_{q,v} \mid z, m, \rho_\eta) \). This last probability will be used to jointly estimate \( x \) and \( m \). The proposed Gibbs sampling algorithm alternatively carries out the following steps:

1. sample \( w \sim p(w \mid y, x, \rho_e) \)
2. sample \( z \sim p(z \mid x, m) \)
3. sample \( (x, m) \sim p(x, m \mid w, z, \rho_\eta, \{ \mu_k^t, \mu_k^t, \sigma_k^t, \sigma_k^t \}_k) \)
4. sample \( \rho_e \sim p(\rho_e \mid y, w, \alpha_e, \beta_e) \)
5. sample \( \rho_\eta \sim p(\rho_\eta \mid w, x, \alpha_\eta, \beta_\eta) \)

As we chose conjugate priors for the variances \( \rho_e \) and \( \rho_\eta \), their \( a \ posteriori \) conditional distribution stay Inverse Gamma whose parameters have only to be updated. There is also no problem for sampling \( z \) in the step 2 because its \( a \ posteriori \) conditional distribution is still a Gibbs field of 4 neighbors. In the following we will develop more in details the steps 1 and 3.

**sampling \( w \)**

Because we assume independent Gaussian noises \( \epsilon_{q,v} \) and independent Gaussian noises \( \eta_{q,v} \) we have from the coupled equations

\[
p(y \mid w) = \prod_{q,v} p(y_{q,v} \mid w_{q,v}) \propto \prod_{q,v} \exp \left\{ -\frac{1}{2\rho_e} ||y_{q,v} - G_q^S w_{q,v}||^2 \right\}
\]

\[
p(w \mid x) = \prod_{q,v} p(w_{q,v} \mid x) \propto \prod_{q,v} \exp \left\{ -\frac{1}{2\rho_\eta} ||w_{q,v} - X_q \phi_{q,v} - X_q G_q D_q w_{q,v}||^2 \right\}
\]

and using the Bayes theorem we note that the probability

\[
p(w \mid y, x, \rho_e, \rho_\eta) \propto \prod_{q,v} p(y_{q,v} \mid w_{q,v}, \rho_e) p(w_{q,v} \mid x, \rho_\eta)
\]
is separable among \( q \) and \( v \), and each \( p(w_{q,v}|y_{q,v},x,\rho_e,\rho_\eta) \) is Gaussian. However the mean and covariance matrix of these Gaussian distributions are intricate to compute due to the size of the matrices \( G^S_{q,v} \) and \( G^D_{q} \). That is why in this step we approximate samples of \( w_{q,v} \) by the samples that maximize the probabilities \( p(w_{q,v}|y_{q,v},x,\rho_e,\rho_\eta) \) obtained by a gradient based optimization algorithm.

**sampling \( (x,m) \)**

The product law first gives

\[
p(x,m|w,z,\rho_\eta,\mu'_k,\mu'_l,\sigma'_k,\sigma'_l) = p(x|m,w,z,\rho_\eta) p(m|w,z,\rho_\eta,\mu'_k,\mu'_l,\sigma'_k,\sigma'_l),
\]

then the sampling of \( (x,m) \) is given by first sampling \( m \) and after \( x \). A sampling following the distribution \( p(m|w,z,\rho_\eta,\mu'_k,\mu'_l,\sigma'_k,\sigma'_l) \) is tractable since the probability \( p(w_{q,v}|z,m,\rho_\eta) \) is separable among the pixels (see appendix). Then we can write for each \( m_k \)

\[
p(m_k|w,z,\rho_\eta,\mu'_k,\mu'_l,\sigma'_k,\sigma'_l) \propto p(m_k|\mu'_k,\mu'_l,\sigma'_k,\sigma'_l) \prod_{q,v} p(w_{q,v}(r)|z(r) = k, m_k, \rho_\eta)
\]

where \( R_k = \{ r; z(r) = k \} \) and with \( p(m_k|\mu'_k,\mu'_l,\sigma'_k,\sigma'_l) \) the Gaussian prior probability of \( m_k \). This conditional a posteriori distribution stay Gaussian and is then easy to sample. For the sampling of \( x \) we still use the Bayes theorem to obtain

\[
p(x|w,z,m,\rho_\eta) \propto p(x|z,m) p(w|x,\rho_\eta) \propto p(x|z,m) \prod_{q,v} p(w_{q,v}|x,\rho_\eta)
\]

This posterior distribution is Gaussian whose covariance matrix is diagonal (see appendix); hence an exact sample of this distribution can be easily obtained.

**SIMULATION RESULTS**

The first test is on simulated data: an object composed of metallic \( (x(r) = 0.5 \) i) and dielectric \( (x(r) = 0.45) \) parts is illuminated by incident fields of 2 frequencies (9 and 10 GHz) and 9 views around it. A Gaussian noise has been added on the measured scattered fields in order to have a signal-to-noise ratio of 10dB. The second test is on experimental data collected in a laboratory-controlled experiment led at the Institut Fresnel (Marseille, France). The scattered fields are here measured from incident fields of two frequencies (4 and 8 GHz) and 18 views around the object. Figure (2) and (3) show the respective results of the proposed algorithm in relation to those obtained by the contrast source inversion (CSI) method. For both methods we observe a good localization of the materials. However we can note two main remarks: the first one is that our proposed method better retrieves the shape of the materials and reconstructs constant regions, even if it seems to overestimate their size. Indeed in the case of experimental data the retrieved dielectric part is 9 cm wide instead of 8 cm in theory, and the retrieved metallic part is 30 mm wide instead of 28 mm in theory. The second remark is also on the experimental data
FIGURE 2. Reconstruction results on simulated data. The test domain is a 17.85 cm sided square partitionned into $51 \times 51$ pixels. (a) real (up) and imaginary (down) of the original object composed of a 35 mm sided dielectric square and a metallic disk with a diameter of 28 mm. (b) reconstruction result with CSI method. (c) reconstruction result with the proposed method.

FIGURE 3. Reconstruction results on experimental data (courtesy of the Institut Fresnel, Marseille, France). The test domain is a 17.85 cm sided square partitionned into $51 \times 51$ pixels. (a) real (up) and imaginary (down) of the original object composed of a dielectric disk with a diameter of 8 cm and a metallic disk with a diameter of 28 mm. (b) reconstruction result with CSI method. (c) reconstruction result with the proposed method.

where the upper part of the dielectric material is not retrieved by any method. This is due to the fact that in theory the electromagnetic wave does not penetrate the metal. This implies that there is a dark zone behind the metal for any wave coming from the top. This also implies an indeterminacy inside the metal. That is why for both (simulated
and experimental) data sets the CSI method reconstructs a highly valued real part in the center of the metallic material to compensate this indeterminacy. The introduction of the knowledge that the materials are quite constant allows not only to reconstruct homogeneous imaginary parts but also to remove high values on the corresponding real part.

**CONCLUSION**

We studied a nonlinear inverse problem in microwave imaging. We proposed a solution to introduce the prior knowledge of the number of materials that compose the unknown object $x$. We used the bilinearity property of the nonlinear model to propose a Gibbs sampling algorithm. A change of variable permitted to obtain separable, and then more tractable, posterior probability distributions and then to jointly sample the object $x$ and the hyper-parameter $m$. Our future studies are on the joint estimation of the three variables $x$, $z$ and $m$ by approximating the conditional a posteriori distribution $p(z|x)$.

**REFERENCES**


**APPENDIX**

In the following the superscripts "r" and "i" will denote respectively the real and imaginary parts of a vector or a matrix. By this way any vector $v = v^r + iv^i \in \mathbb{C}^{ND}$ can be written in $\mathbb{R}^{2ND}$ by $v = [v^r \ v^i]^T$, and any $(ND \times ND)$ matrix $D = D^r + i D^i$ can be written by a $(2ND \times 2ND)$ matrix $D = \begin{bmatrix} D^r & -D^i \\ D^i & D^r \end{bmatrix}$. If we also note $\phi_{q,v} = \text{diag}(\phi_{0,q,v} + G_q^D w_{q,v}) = \phi_{0,q,v}^r + i \phi_{0,q,v}^i$, it comes that the state equation (6) can be rewritten in $\mathbb{R}^{2ND}$:

$$
\begin{bmatrix}
w_{q,v}^r \\
w_{q,v}^i
\end{bmatrix} = \Phi_{q,v} F_q \begin{bmatrix} x^r \\ x^i \end{bmatrix} + \eta_{q,v}
$$

(14)
where \( \Phi_{q,v} = \begin{bmatrix} \phi_{q,v}^r & -\phi_{q,v}^i \\ \phi_{q,v}^i & \phi_{q,v}^r \end{bmatrix} \) and \( F_q = \begin{bmatrix} I_{Nd} & 0 \\ 0 & \frac{1}{\Omega_q} I_{Nd} \end{bmatrix} \). We can also rewrite the prior probability of \( x \) given \( z \) and \( m \) in \( \mathbb{R}^{2Nd} \):

\[
p(x|z, m) = \mathcal{N}(m_z, \Sigma_z)
\]

where \( m_z(r) = m_k^r, m_z(r + Nd) = m_k^i \), \( \Sigma_z(r, r) = \rho_k^r \) and \( \Sigma_z(r + Nd, r + Nd) = \rho_k^i \) if \( z(r) = k \), for any \( r = 1, \ldots, Nd \).

### A. Complex separability of \( p(w_{q,v}|z, m, \rho_\eta) \)

From the relations (14) and (15) we deduce

\[
p(w_{q,v}|z, m, \rho_\eta) = \mathcal{N}(\Phi_{q,v} F_q m_z, \Phi_{q,v} F_q \Sigma_z F_q^T \Phi_{q,v}^T + \rho_\eta I_{Nd})
\]

The covariance matrix of this Gaussian distribution can be written by blocks

\[
\Phi_{q,v} F_q \Sigma_z F_q^T \Phi_{q,v}^T + \rho_\eta I_{Nd} = \begin{bmatrix} R_1 & R_3 \\ R_3^T & R_2 \end{bmatrix},
\]

with

\[
R_1 = \phi_q^r \Sigma_z \phi_q^r + \frac{1}{\Omega_q} \phi_q^i \Sigma_z \phi_q^i + \rho_\eta I_{Nd}
\]

\[
R_2 = \phi_q^r \Sigma_z \phi_q^i + \frac{1}{\Omega_q} \phi_q^r \Sigma_z \phi_q^i + \rho_\eta I_{Nd}
\]

\[
R_3 = \phi_q^i \Sigma_z \phi_q^i - \frac{1}{\Omega_q} \phi_q^r \Sigma_z \phi_q^i
\]

and \( R_1, R_2, R_3 \) are diagonal \( Nd \times Nd \) matrices. This implies that there is only a correlation between \( w_{q,v}(r) \) and \( w_{q,v}(r + Nd), \forall r = 1, \ldots, Nd \), i.e. between the real part \( w_{q,v}^r(r) \) and the imaginary part \( w_{q,v}^i(r) \) of a pixel \( w_{q,v}(r) \in \mathbb{C} \). Hence the probability distribution is separable:

\[
p(w_{q,v}|z, m, \rho_\eta) = \prod_{r \in D} p(w_{q,v}(r)|z(r) = k, m_k, \rho_\eta)
\]

where \( p(w_{q,v}(r)|z(r) = k, m_k, \rho_\eta) \) is the probability distribution of the complex variable \( w_{q,v}(r) \).

### B. Posterior distribution of \( x \)

Using the relations (13) and (15) we deduce that the posterior distribution of \( x \) is still Gaussian \( p(x|w, z, m, \rho_\eta) = \mathcal{N}(m_x^{apost}, \Sigma_x^{apost}) \), with

\[
m_x^{apost} = \Sigma_x^{apost} \left( \begin{bmatrix} \frac{1}{\rho_\eta} \mathcal{R} \mathcal{E} \sum_{q,v} \phi_{q,v}^* w_{q,v} \\ \frac{1}{\rho_\eta} \mathcal{I} m \sum_{q,v} \frac{1}{\Omega_q} \phi_{q,v}^* w_{q,v} \end{bmatrix} + \Sigma_z^{-1} m_z \right)
\]

\[
\Sigma_x^{apost} = \begin{bmatrix} \frac{1}{\rho_\eta} \sum_{q,v} \phi_{q,v}^* \phi_{q,v} & 0 \\ 0 & \frac{1}{\rho_\eta} \sum_{q,v} \frac{1}{\Omega_q} \phi_{q,v}^* \phi_{q,v} \end{bmatrix} + \Sigma_z^{-1}.
\]