A BAYESIAN APPROACH FOR NONLINEAR INVERSE SCATTERING TOMOGRAPHIC IMAGING

Hervé Carfantan and Ali Mohammad-Djafari

Laboratoire des Signaux et Systèmes (CNRS-ESE-UPS)
Plateau de Moulon, 91192 Gif-sur-Yvette Cedex, France.
carfantan@lss.supelec.fr, djafari@lss.supelec.fr

ABSTRACT

We propose a new method to solve the nonlinear inverse problem of tomographic imaging using microwave or ultrasound probing, beyond the classical first order Born or Rytov approximations. The relation between the data and the measurement is given by two coupled nonlinear equations.

We set this problem as one of estimation and propose a solution within the Bayesian probability framework. The Maximum *A Posteriori* estimate determination leads to a multi-modal criterion minimisation. Global minimisation using Simulated Annealing is not practicable due to the high calculation cost. We propose a feasible deterministic relaxation algorithm inspired by the Graduated Non Convexity principle to perform this minimisation.

1. INTRODUCTION

Tomographic imaging with scattering waves such as ultrasound or microwave arises in various areas such as medical imaging, non-destructive testing and geophysical remote sensing. The purpose is to construct an image representing some physical properties of an object from a set of measured field data. The image - data relation is not linear and is given by two coupled equations. It is linearised customarily using the Born or Rytov approximations. These approximations, however, break down when the object to be reconstructed is too large or has a too high contrast.

Several authors have tried to circumvent the nonlinearity by solving iteratively and independently the two coupled linear equations [1], others have proposed methods for solving these equations by defining and minimising an intuitive criterion [2].

In our work, we try to find a regularised solution of this nonlinear ill-posed inverse problem within a Bayesian probability framework. We define the solution as the Maximum *A Posteriori* (MAP) estimate, which takes into account the measurement noise, supposed to be white and Gaussian, and the correlated nature of the image.

The MAP estimate computation requires the minimisation of a criterion which may have local minima. Techniques such as Simulated Annealing (SA) can not be used in our case due to their high calculation cost. This is due to the fact that the relation between the data and the object is not local.

We propose a deterministic relaxation algorithm in order to perform this minimisation. We have no theoretical guarantee of convergence of this algorithm, but it has been shown to give satisfactory experimental results.

2. MODELLING AND BACKGROUND

The two-dimensional diffraction tomography or inverse scattering problem is a classical one, and the forward modelling equations and their establishment can be found in many text books and other works. See for example [1]. The geometrical configuration of the problem is shown in Fig. 1. We want to reconstruct the complex permittivity profile x of an object from the observation of the scattered field y. From Maxwell equations we can derive the following coupled equations relating x and y in an operator form:

$$y = G_{\rm m}(x, \phi), \quad (1)$$
$$\phi = \phi^{\rm inc} + G_{\rm o}(x, \phi), \quad (2)$$

where $\phi^{\rm inc}$ and ϕ are the incident and the total field on the object, and $G_{\rm m}$, $G_{\rm o}$ are linear operators with respect to x and ϕ, related to the Green functions.

The classical Born approximation neglects the scattered field on the object in Eq. (2) which amounts to $G_{\rm o} = 0$. This leads to the linear equation relating x and y:

$$y = G_{\rm m}\left(x, \phi^{\rm inc}\right).$$

This linear inverse problem has preoccupied many researchers in the past and different solutions, for most

0-7803-2431-5/95 $4.00 © 1995 IEEE

favourable experimental conditions, can be found in the literature. However, this approximation is too restrictive, and several authors have recently tried to go beyond this limitation.

The Born Iterative Method (BIM) [1] (or related methods) treats this nonlinear problem by solving iteratively (once discretised) the Eq. (1) with respect to x for fixed ϕ_n and the Eq. (2) with respect to ϕ for fixed x_n. This leads to solve successively linear inverse problems. The limits of convergence of such an algorithm are still to be specified. It has given satisfactory results in some experiments, but failed in others.

R.E. Kleinman [2] proposed recently a new method. He established intuitively a quadratic criterion, related to the errors on the equations (1) and (2):

$$J_{\mathrm{K}}(x,\phi)=\frac{\|y-G_{\mathrm{m}}(x,\phi)\|^2}{\|y\|^2}+\frac{\|\phi-\phi^{\mathrm{inc}}-G_{\mathrm{o}}(x,\phi)\|^2}{\|\phi^{\mathrm{inc}}\|^2}.$$

Then he tried to minimise it simultaneously with respect to x and ϕ using a gradient descent type algorithm. The solution found with this algorithm can be a local minimum of the given criterion.

All these methods give some good practical results as long as the contrast is not too high and as long as the signal to noise ratio (SNR) is not too low.

As an alternative, we propose to define an objective criterion to minimise, which takes into account the modelling and measurement noise and the prior information on the image to reconstruct.

3. PROPOSED METHOD

3.1. Definition of an objective criterion

The discretisation of the Eq. (1-2) gives, in a compact notation:

$$\begin{aligned} \boldsymbol{y} &= \boldsymbol{G}_{\mathrm{m}} X \boldsymbol{\phi}, \\ \boldsymbol{\phi} &= \boldsymbol{\phi}^{\mathrm{inc}} + \boldsymbol{G}_{\mathrm{o}} X \boldsymbol{\phi}, \end{aligned}$$

where $\boldsymbol{G}_{\mathrm{m}}, \boldsymbol{G}_{\mathrm{o}}$ are matrices related to the Green functions, X is a diagonal matrix with the components of the vector $\boldsymbol{x}$ (a n length vector) as its diagonal elements and $\boldsymbol{y}, \boldsymbol{\phi}$ are respectively m and n length vectors representing the measured data and the total wave components on the object ($n > m$).

These two equations can be combined to obtain a symbolic explicit relation between the data $\boldsymbol{y}$ and the unknowns $\boldsymbol{x}$:

$$\boldsymbol{y}=\boldsymbol{G}_{\mathrm{m}} X\left(I-\boldsymbol{G}_{\mathrm{o}} X\right)^{-1} \boldsymbol{\phi}^{\mathrm{inc}}=\boldsymbol{A}(\boldsymbol{x}),$$

where the considered matrix is supposed to be invertible.

Finding $\boldsymbol{x}$ for given data $\boldsymbol{y}$ is an inherently ill-posed problem and its resolution requires prior information on the solution and on the errors.

In a probabilistic framework, to take into account the unavoidable uncertainties on the data, we consider the equation:

$$\boldsymbol{y}=\boldsymbol{A}(\boldsymbol{x})+\boldsymbol{b},$$

where $\boldsymbol{b}$ stands for the errors (modelling and discretisation errors as well as the noise), which is considered to be Gaussian $\boldsymbol{b} \sim N(\mathbf{0}, \sigma_b^2 I)$, with a known variance σ_b^2. From this assumption we deduce the conditional law:

$$p(\boldsymbol{y}|\boldsymbol{x})=\left(\frac{1}{\sqrt{2\pi}\sigma_b}\right)^m \exp\left(-\frac{1}{2\sigma_b^2}\|\boldsymbol{y}-\boldsymbol{A}(\boldsymbol{x})\|^2\right).$$

The Bayesian approach gives us a coherent setting to introduce the information on the solution by defining a prior probability law $p(\boldsymbol{x})$ [3]. For feasibility requirements we take a prior law of the form:

$$p(\boldsymbol{x})=\frac{1}{Z}\exp\left[-\mu U(\boldsymbol{x})\right],$$

with Z a normalising factor, and U an energy function. Then, Bayes' rule allows to exhibit the posterior law:

$$p(\boldsymbol{x}|\boldsymbol{y}) \propto p(\boldsymbol{y}|\boldsymbol{x})p(\boldsymbol{x}).$$

We then define the solution to be the Maximum *A Posteriori* estimate of $\boldsymbol{x}$:

$$\hat{\boldsymbol{x}}=\arg\max_{\boldsymbol{x}}\{p(\boldsymbol{x}|\boldsymbol{y})\}=\arg\min_{\boldsymbol{x}}\{J(\boldsymbol{x})\},$$

with

$$J(\boldsymbol{x})=\frac{1}{2\sigma_b^2}\|\boldsymbol{y}-\boldsymbol{A}(\boldsymbol{x})\|^2+\mu U(\boldsymbol{x}),$$

where μ plays the role of a regularisation parameter.

The choice of the energy function U is essential in the Bayesian framework. It has to incorporate our knowledge on the object to reconstruct. A very powerful class of models for images is the Markov Random Fields (MRF) [4], its goal is to take into account the correlated nature of the image to reconstruct. Gibbs distributions are used to explicitly write MRF's distributions. Their energy functions could be written:

$$U(\boldsymbol{x})=\sum_c V_c(\boldsymbol{x}),$$

where V_c is a function (called potential) of local groups of points c (cliques).

In this work, we only took into account potential functions of the form:

$$U(\boldsymbol{x})=\sum_{\{i,j \text{ neighbours}\}}|x_j-x_i|^p,\ 1<p\leq 2,$$

which restricts us to a class of MRFs called Generalised Gauss Markov Random Fields [5], and horizontal and vertical neighbourhood (first order). For p near or equal 2, it severely penalises large differences between the values of the neighbours and is useful to reconstruct smooth objects, on the other hand, for p near to 1, it allows such differences, and could be used to reconstruct discontinuous objects.

3.2. Proposed optimisation method

The criterion $J(\boldsymbol{x})$ may have many local minima. It's minimisation should be done with a global optimisation algorithm.

Simulated annealing is a global optimisation technique which has been used in image restoration [4]. It updates successively the pixel values of the image according to the conditional posterior probability law. It is particularly efficient when this distribution has local properties, which makes the cost of calculation very low for each update. However, in our problem, due to the large support of the operator $\boldsymbol{A}$, the SA is practically inextricable. In fact, the posterior distribution is still a Gibbs distribution but the neighbourhood of a pixel becomes practically the whole image.

We introduce a deterministic relaxation algorithm inspired by the Graduated Non Convexity (GNC) principle proposed by Blake and Zisserman in [6] for noise cancellation and segmentation and extended to the general linear ill-posed inverse problem in [7]. The principle of this algorithm is very simple. It consists in approximating the non convex criterion $J(\boldsymbol{x})$ with a sequence of continuously derivable criteria $J_n(\boldsymbol{x})$ converging toward it for $n \to \infty$, by taking care to choose the first one $J_0(\boldsymbol{x})$ to be convex. Then we minimise each criterion $J_n(\boldsymbol{x})$, with a local optimisation technique such as adaptive gradient, using as starting point the minimum of the precedent criterion.

Up to now, such a principle has only been studied for linear inverse problems where the log-likelihood is a convex function and the multi-modality of the MAP criterion is due to the prior law part. In this work, we consider prior laws such that $U(\boldsymbol{x})$ is convex, so that the multi-modality of $J(\boldsymbol{x})$ is due to the non-convexity of the log-likelihood, which comes from the nonlinearity of the operator $\boldsymbol{A}(\boldsymbol{x})$.

To introduce the GNC technique, we consider the sequence:

$$\boldsymbol{A}_n(\boldsymbol{x}) = \boldsymbol{G}_{\mathrm{m}} X (I - r_n \boldsymbol{G}_{\mathrm{o}} X)^{-1} \boldsymbol{\phi}^{\mathrm{inc}},$$

with $r_0 = 0$, and $\lim_{n\to\infty} r_n = 1$. Note that the first term ($r_0 = 0$) corresponds to the case of Born approximation with a linear relation between the data and the unknowns, which results to the following convex criterion:

$$J_0(\boldsymbol{x}) = \frac{1}{2\sigma_b^2} \|\boldsymbol{y} - \boldsymbol{G}_{\mathrm{m}} X \boldsymbol{\phi}^{\mathrm{inc}}\|^2 + \mu U(\boldsymbol{x}).$$

The sequence $J_n(\boldsymbol{x})$ corresponding to $\boldsymbol{A}_n(\boldsymbol{x})$ converges toward $J_n(\boldsymbol{x})$ for $n \to \infty$.

4. SIMULATIONS AND RESULTS

The main objectives of these simulations were to show that:

- The criterion may have many local minima. This has been shown empirically by choosing many arbitrary initialisations and noting that a gradient descent technique gives notably different solutions.
- The proposed GNC based optimisation algorithm permits always to reach the same solution, for any arbitrary initialisation, and we can hopefully assume this solution to be a global minimum.
- The solution is not very sensitive to the choice of the relaxation scheme r_n and to the parameter μ. In practical experiments, a low number of relaxation steps has been sufficient. For example $r_n = \frac{n}{10}, n = 0 \ldots 10$.
- The obtained results are more satisfactory than the solutions obtained by the BIM or by Kleinman's method. In fact, in some cases where the contrast is too high, the last two methods do not give satisfactory results.

In the following figures we give examples of obtained results. The simulated object has a diameter of one wavelength, discretised in 11×11 points. There are 6 sources-receivers equally spaced around the object which gives 36 measurement data. The SNR is of 20dB. We present the results given by the proposed method, Born iterative and Kleinman's method for a smooth object with relative peak contrast of 3, and the results given by the proposed method and BIM for a non smooth object with relative peak contrast of 2.

Our method gives better results than the others and permits to take into account the discontinuity with better efficiency than the BIM, even with the same type of regularisation, while Kleinman's method can not take them into account.

5. CONCLUSION

We proposed a new method, based on the Bayesian estimation framework, to solve the classical scattering waves diffraction tomography imaging. The main advantage of an inversion method based on this approach is to be able to define an objective criterion to optimise and tools to characterise the obtained solution.

Furthermore, it gives a powerful tool to introduce prior information on the image to reconstruct.

The MAP estimate computation leads to a multimodal criterion. Its minimisation is performed with a new deterministic relaxation method, based on the GNC principle, which gives good practical results.

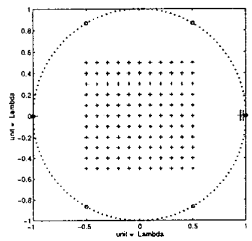

Figure 1: Geometrical configuration.

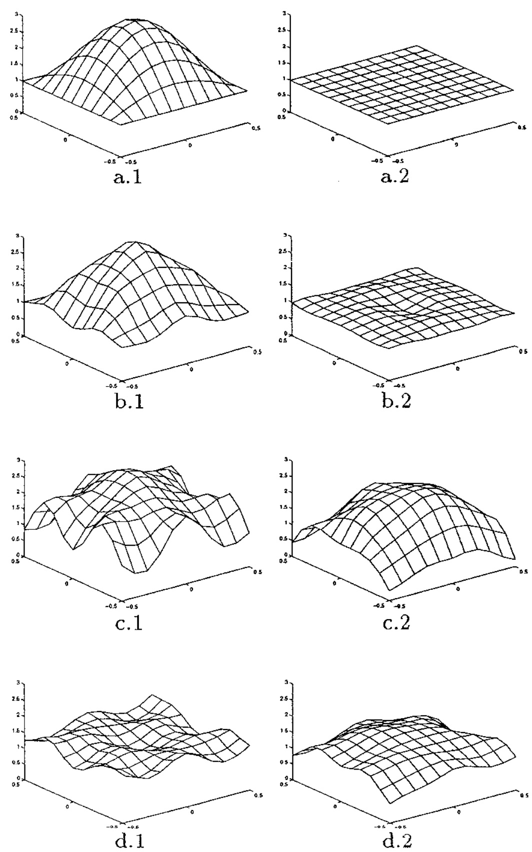

Figure 2: Reconstruction results for a smooth object with a maximum contrast peak of 3.
a) Original contrast: real and imaginary part,
b) Proposed method with $p = 2$, $D = 0.02$,
c) Kleinman's method $D = 0.84$,
d) Born Iterative Method $D = 0.78$,
D is the relative quadratic distance of the reconstruction to the original, $D = \frac{\|\hat{x} - x\|^2}{\|x\|^2}$.

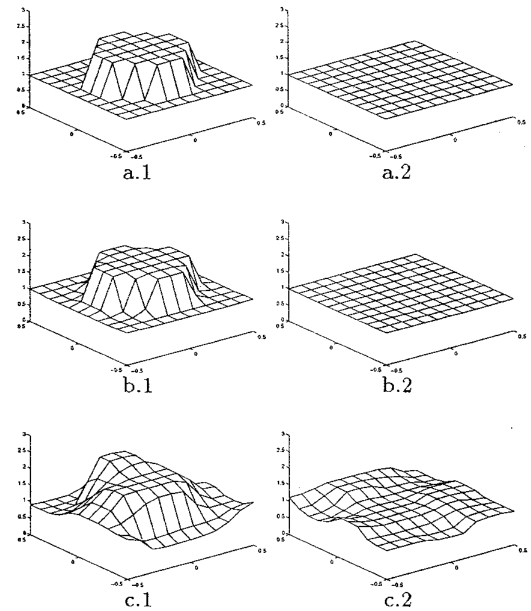

Figure 3: Reconstruction results for a non smooth object with a maximum peak contrast of 2.
a) Original contrast: real and imaginary part,
b) Proposed method with $p = 1.1$,
c) Born Iterative Method with same type of regularization.

6. REFERENCES

[1] Y. M. Wang, and W. C. Chew, "An iterative solution of the two-dimensional electro-magnetic inverse scattering problem," *Int. Journal of Imaging Systems and Technology,* vol. 1, p. 100-108 (1989).

[2] R. E. Kleinman, and P. M. van den Berg, "A modified gradient method for two dimensional problems in tomography," *Journal of Computational and Applied Mathematics,* vol. 42, p. 17-35 (1992).

[3] G. Demoment, "Image reconstruction and restoration: overview of common estimation structures and problems," *IEEE Trans. Acoustic, Speech, ans Signal Processing,* vol. 37, p. 2024-2036 (1989).

[4] S. Geman and D. Geman "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Patt. Anal. Mach. Intell.,* vol. 6, pp 721-741, (1984).

[5] C. Bouman and K. Sauer "A generalized Gaussian image model for edge-preserving MAP estimation," *IEEE Trans. Image Processing,* vol. 2, pp 296-310, (1993).

[6] A. Blake, and A. Zisserman, "Visual reconstruction," *MIT Press, Cambridge MA,* (1987).

[7] M. Nikolova, A. Mohammad-Djafari, and J. Idier, "Inversion of large-support ill-conditioned linear operators using a Markov model with a line process," *IEEE ICASSP 1994,* vol. 5, p. 357-360 (1994).