A SINGLE SITE UPDATE ALGORITHM FOR NONLINEAR DIFFRACTION TOMOGRAPHY

H. Carfantan, A. Mohammad-Djafari and J. Idier

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, idier@lss.supelec.fr

ABSTRACT

We focus on the nonlinear inverse problem of diffraction tomography. We set the problem as one of estimation within the Bayesian framework and define the solution as the maximum *a posteriori* (MAP) estimate which corresponds to the global minimum of a multimodal criterion. The objective of this paper is to present a new deterministic single site update algorithm specially designed to compute this solution. The term of fidelity to the data, function of one pixel value, can be written as a second order rational fraction. Thus, the 1-D MAP criterion can be evaluated – and minimized – at a very low computational cost. Moreover, for certain MRF models the global minimum can even be computed explicitly as roots of a polynomial. The proposed algorithm turns these properties to advantage and moreover performs the updates of intermediate quantities at a particularly low cost compared to the criterion evaluation. Even if not guaranteed to converge towards the global minimum, the algorithm has shown itself to give satisfactory practical results.

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 the spatial variation of some physical properties of an object from a finite set of field data scattered by this object. The object-data relation is not linear and is given by two coupled equations. It is linearized customarily using the Born or Rytov approximations leading to the well known linear diffraction tomography. These approximations, however, break down when the object to be reconstructed is too large or has a too high contrast [1].

We already considered this nonlinear inverse problem within the Bayesian estimation framework [2]. We defined a regularized solution as the Maximum *a posteriori* (MAP) estimate which account for both errors on the measurement – assumed centered white Gaussian circular – and prior knowledge on the object – *e.g.* modeled with Markov random fields. The MAP estimate computation requires the minimization of a criterion which may have local minima.

This standpoint allowed us to establish links between some existing methods as algorithms to compute the regularized solution [2]. Most of these methods are based on local optimization techniques and can easily get stuck in local minima in difficult situations.

We also proposed a deterministic relaxation algorithm based on the graduated non convexity (GNC) principle in order to perform this global minimization [3]. Even if there is no guarantee to reach the global minimum with such an algorithm, it has given very satisfactory results in these difficult situations compared to other local minimization techniques. However, its computational cost is very expensive.

Different algorithms in the class of Single Site Updates Algorithms (SSUA) have been studied in image processing. Deterministic [4] and stochastic [5] ones have given very interesting results [6]. In this paper, we propose a deterministic SSUA specially designed to solve the nonlinear diffraction tomography inverse problem. This algorithm takes benefit of the particular form of the forward model (thus of the likelihood function) and possibly of the prior model, to perform a global 1-D optimization for each pixel. Of course, it is a sub-optimal minimization technique which can get stuck in a local minimum, however, it has given satisfactory results for a relative cheap computational cost and with a high convergence rate.

2. MODELING AND BACKGROUND

The geometrical configuration of the problem is shown in Fig. 1. The objective is to reconstruct the 2-D complex permittivity profile $x(\boldsymbol{r}), \boldsymbol{r} \in D_{\mathrm{O}}$ of a cylindrical object from the observation of the scattered field $y(\boldsymbol{r}_j), \boldsymbol{r}_j \in D_{\mathrm{M}}$ measured at n_{M} sensors for n_{S} distinct wave emission.

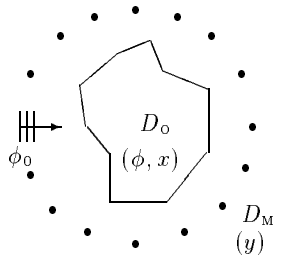

Figure 1: 2-D cylindrical geometrical configuration.

From Maxwell equations, one can derive coupled equations relating x and y in an operator form:

$$y = \mathcal{G}_{\mathrm{M}}(x, \phi), \tag{1a}$$

$$\phi = \phi_0 + \mathcal{G}_{\mathrm{O}}(x, \phi), \tag{1b}$$

where ϕ_0 and ϕ are the incident and the total field on the object, and $\mathcal{G}_\text{M}$, $\mathcal{G}_\text{O}$ are bilinear operators with respect to x and ϕ, related to the Green function of the homogeneous background. The classical Born approximation neglects the scattered field on the object in Eq. (1b) – which amounts to $\phi = \phi_0$ – and leads to a linear relation between x and y. However, this approximation is too restrictive [1], and a recent challenge is to try to go beyond its limitations. Once the Eq. (1a–1b) are discretized, we can derive the following explicit relation between the unknown $\boldsymbol{x} \in \mathbb{C}^{n_\text{O}}$ and the data $\boldsymbol{y} \in \mathbb{C}^{n_\text{M} \times n_\text{S}}$: $\boldsymbol{y} = \mathcal{A}(\boldsymbol{x})$, with

$$\mathcal{A}(\boldsymbol{x}) = \boldsymbol{G}_\text{M} \boldsymbol{X} \left(\boldsymbol{I} - \boldsymbol{G}_\text{O} \boldsymbol{X}\right)^{-1} \boldsymbol{\phi}_0,$$

where $\boldsymbol{G}_\text{M}$, $\boldsymbol{G}_\text{O}$ are matrices and $\boldsymbol{X}$ is a diagonal matrix with the components of the vector $\boldsymbol{x}$ as its diagonal elements. Note that $\mathcal{A}(\boldsymbol{x})$ is nonlinear and of large support.

To take into account the unavoidable uncertainties on the data, we consider the equation:

$$\boldsymbol{y} = \mathcal{A}(\boldsymbol{x}) + \boldsymbol{n}, \tag{2}$$

where $\boldsymbol{n}$ stands for the errors on the measurement, modeled with a circular Gaussian noise $\boldsymbol{n} \sim N(\boldsymbol{0}, \sigma_n^2 \boldsymbol{I})$, with a known σ_n^2 and independent of $\boldsymbol{x}$.

We then define the solution as the MAP estimate of $\boldsymbol{x}$:

$$\boldsymbol{x}_\text{MAP} = \arg\max_{\boldsymbol{x}} p(\boldsymbol{x} \mid \boldsymbol{y}) = \arg\min_{\boldsymbol{x}} \mathcal{J}_\text{MAP}(\boldsymbol{x})$$

with

$$\mathcal{J}_\text{MAP}(\boldsymbol{x}) = \mathcal{J}_\text{ML}(\boldsymbol{x}) + \lambda \mathcal{U}(\boldsymbol{x}).$$

In this criterion, $\mathcal{J}_\text{ML}$ (data part) is the criterion to be minimized when the maximum likelihood (ML) estimate is considered:

$$\mathcal{J}_\text{ML}(\boldsymbol{x}) = \|\boldsymbol{y} - \mathcal{A}(\boldsymbol{x})\|^2,$$

$\mathcal{U}$ is the prior energy[1] (regularization term) and λ plays the role of a regularization parameter.

We chose to model the *a priori* information with Markov random fields (MRF) which are classically used in image processing. The major difficulty of the problem is due to the non-linearity of the forward model (operator $\mathcal{A}$) which implies the non-convexity of the data part, so we restricted ourself to MRFs with convex energy functions.

Even with a convex energy function $\mathcal{U}$, the criterion may have local minima due to the non-linearity of $\mathcal{A}$. Note that for easy configurations (large number of data and low contrast), where there is no local minima, local minimization techniques may be sufficient. In fact, a great number of methods are available in this case to compute $\boldsymbol{x}_\text{MAP}$ with a lower cost than the direct minimization of $\mathcal{J}_\text{MAP}$ [2]. However, in difficult configurations, they can get stuck in local minima and global optimization techniques may be used.

Stochastic algorithms such as SA or MCMC are practically inextricable because of the large support and computational cost of the operator $\mathcal{A}$. The global deterministic algorithm proposed in [3] has given satisfactory results but with a still high computational cost. It is based on local minimization of a sequence of criteria of the same form than $\mathcal{J}_\text{MAP}$, globally approximating $\mathcal{J}_\text{MAP}$, with a first order descent technique. Note that one iteration of a first order descent technique requires the computation of the gradient of the criterion and at least one evaluation of the criteria depending on the 1-D minimization strategy (in general more than 3 evaluations).

[1] Note that $\boldsymbol{x}$ is complex but it real and imaginary parts have physically distinct meaning. In this paper, we consider energy functions applied independently to each part, but correlation could be introduced.

3. PROPOSED ALGORITHM

Following the work of [4] for the minimization of convex MAP criterion in Bayesian X-ray computed tomography, we propose to minimize this nonconvex criterion updating a single pixel value, at each iteration, rather than the entire image.

We show [7] that the criterion $\mathcal{J}_\text{ML}$, function of a pixel[2] i can be written as a second order rational fraction[3]:

$$\mathcal{J}_\text{ML}(\boldsymbol{x} + \boldsymbol{x}_i) = \gamma + \frac{\alpha_1 x_i + \alpha_2 x_i^2}{\beta_0 + \beta_1 x_i + \beta_2 x_i^2}.$$

An example of this criterion is illustrated Fig. 2. This important feature can be taken into account in such a SSUA, because the criterion evaluation can be made at a very low computational cost, thus the 1-D minimization could be very cheap.

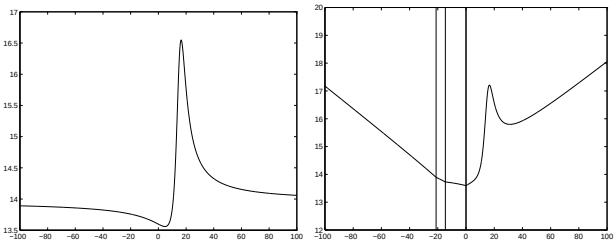

Figure 2: Left: Illustration of the data part of the 1-D criterion: $\mathcal{J}_\text{ML}(\boldsymbol{x} + \boldsymbol{x}_i)$ (rational fraction). Right: Illustration of the 1-D criterion $\mathcal{J}_\text{MAP}(\boldsymbol{x} + \boldsymbol{x}_i)$ for L_1 regularization (piecewise rational fraction).

Moreover, if the energy function $\mathcal{U}$ derives from a Markovian model of potential function $\mathcal{V}_c$, only the potentials of the cliques containing the pixel i are to be changed:

$$\mathcal{U}_i(x_i) \triangleq \mathcal{U}(\boldsymbol{x} + \boldsymbol{x}_i) = \mathcal{U}(\boldsymbol{x}) + \sum_{c, i \in c} \left(\mathcal{V}_c(\boldsymbol{x} + \boldsymbol{x}_i) - \mathcal{V}_c(\boldsymbol{x})\right).$$

So the criterion $\mathcal{J}_i(x_i) \triangleq \mathcal{J}_\text{MAP}(\boldsymbol{x} + \boldsymbol{x}_i)$ can be evaluated at a very low computational cost.

Once determined the new value of the pixel i, different quantities have to be updated to account for this change. Using the matrix inversion lemma applied on

$$\left((\boldsymbol{I} - \boldsymbol{G}_\text{O} \boldsymbol{X}) - \boldsymbol{G}_\text{O} \boldsymbol{X}_i\right)^{-1}$$

[2] In fact function of it real or imaginary part.

[3] In this paper, we respect the following notation: x_i value to add to the current value of pixel i ; $\boldsymbol{1}_i = [0, \ldots, 0, 1, 0, \ldots, 0]^\text{t}$ column vector zero everywhere except at the ith line ; $\boldsymbol{x}_i = x_i \boldsymbol{1}_i$, ; $\mathbb{1}_i$ diagonal matrix zero everywhere except at the ith line, ith column (value 1) ; $\boldsymbol{X}_i = x_i \mathbb{1}_i$; $\boldsymbol{M}^i$ column vector corresponding to the ith column of matrix $\boldsymbol{M}$; $\boldsymbol{M}_i$ line vector corresponding to the ith line of matrix $\boldsymbol{M}$.

which intervenes in $\mathcal{A}(\boldsymbol{x}+\boldsymbol{x}_i)$, these updates can be performed with a relative low cost. Finally, The sweep of all the pixels is performed at a cost of the same order as the gradient of the criterion.

Considering the quantities $\boldsymbol{T}=(\boldsymbol{I}-\boldsymbol{G}_{\mathrm{O}}\boldsymbol{X})^{-1}\boldsymbol{G}_{\mathrm{O}}$ of size $n_{\mathrm{O}}\times n_{\mathrm{O}}$ and $\boldsymbol{R}=\boldsymbol{G}_{\mathrm{M}}\boldsymbol{X}\boldsymbol{T}$ of size $n_{\mathrm{M}}\times n_{\mathrm{O}}$, the algorithm of Table 1 can be deduced. The number of complex multiplications for one pixel is:

$$n_{\mathrm{O}}{}^2+n_{\mathrm{O}}(n_{\mathrm{M}}+n_{\mathrm{S}}+1)+5n_{\mathrm{M}}*n_{\mathrm{S}}+4n_{\mathrm{M}}+n_{\mathrm{S}}.$$

(1-D global minimization excepted).

Table 1: Proposed algorithm

Initialization (corresponding to $\boldsymbol{x}=\boldsymbol{0}$):
$\boldsymbol{T}=\boldsymbol{G}_{\mathrm{O}}$, $\boldsymbol{\phi}=\boldsymbol{\phi}_0$, $\boldsymbol{R}=\boldsymbol{0}$, $\boldsymbol{S}=\boldsymbol{y}$.

1. Selection of a pixel i.

2. Computation of $\boldsymbol{P}$ and ρ :
   $\boldsymbol{P}=(\boldsymbol{G}_{\mathrm{M}}{}^i+\boldsymbol{R}^i)\boldsymbol{\phi}_i$,
   $\rho=\mathbf{1}_i^{\mathrm{t}}(\boldsymbol{I}-\boldsymbol{G}_{\mathrm{O}}\boldsymbol{X})^{-1}\boldsymbol{G}_{\mathrm{O}}{}^i=\boldsymbol{T}_i^i$.

3. Computation of coefficients α_k and β_k
   For the real part of i: $\mathcal{R}\{x_i\}$

$\gamma=\boldsymbol{S}^\dagger\boldsymbol{S}$	$\beta_0=1$
$\alpha_1=-2\mathcal{R}\{\boldsymbol{S}^\dagger\boldsymbol{P}\}$	$\beta_1=-2\mathcal{R}\{\rho\}$
$\alpha_2=\boldsymbol{P}^\dagger\boldsymbol{P}+2\mathcal{R}\{\rho\boldsymbol{P}^\dagger\boldsymbol{S}\}$	$\beta_2=\lvert\rho\rvert^2$

   For the imaginary part of i: $\mathcal{I}\{x_i\}$

$\gamma=\boldsymbol{S}^\dagger\boldsymbol{S}$	$\beta_0=1$
$\alpha_1=2\mathcal{I}\{\boldsymbol{S}^\dagger\boldsymbol{P}\}$	$\beta_1=2\mathcal{I}\{\rho\}$
$\alpha_2=\boldsymbol{P}^\dagger\boldsymbol{P}+2\mathcal{R}\{\rho\boldsymbol{P}^\dagger\boldsymbol{S}\}$	$\beta_2=\lvert\rho\rvert^2$

4. Computation of the new value for the real part or the imaginary part of pixel i minimizing $\mathcal{J}_i(x_i)$: 1-D global minimization.

5. Updates of $\boldsymbol{x}=\boldsymbol{x}+x_i\mathbf{1}_i$, $\xi=\frac{x_i}{1-\rho x_i}$
   and of matrices $\boldsymbol{S}$, $\boldsymbol{R}$, $\boldsymbol{\phi}$ and $\boldsymbol{T}$ (in this order):
   $\boldsymbol{S}=\boldsymbol{S}-\left(\xi\boldsymbol{R}^i+x_i(1+\xi\boldsymbol{T}_i^i)\boldsymbol{G}_{\mathrm{M}}{}^i\right)\boldsymbol{\phi}_i$
   $\boldsymbol{R}=\boldsymbol{R}+\left(\xi\boldsymbol{R}^i+x_i(1+\xi\boldsymbol{T}_i^i)\boldsymbol{G}_{\mathrm{M}}{}^i\right)\boldsymbol{T}_i$
   $\boldsymbol{\phi}=\boldsymbol{\phi}+\xi\boldsymbol{T}^i\boldsymbol{\phi}_i$
   $\boldsymbol{T}=\boldsymbol{T}+\xi\boldsymbol{T}^i\boldsymbol{T}_i$

4. Iterate on 1. to 5. until convergence.

4. 1-D GLOBAL MINIMIZATION

At each iteration, the proposed algorithm requires the minimization of $\mathcal{J}_i$ with respect to the real or imaginary part of x_i. If we consider potential functions $\mathcal{V}(t)$ applying on the *transition* between two neighboor pixels; then, if the derivative of $\mathcal{V}$ is a rational fraction, the derivative of the criterion $\mathcal{J}_i(x_i)$ *w.r.t.* x_i is a rational fraction too. Thus, its local extrema correspond to the roots of polynomials and the global minimum can be computed easily.

Among the potential functions proposed in the literature, some of them verify this property. In particular, the well known L_2 and L_1 cases (resp. $\mathcal{V}(t)=t^2$ and $\mathcal{V}(t)=|t|$) but also the case of the Huber function: $\mathcal{V}^T(t)=\frac{|t|^2}{T^2}$ if $|t|\leq T$, $2\frac{|t|}{T}-1$ if $|t|>T$.

4.1. L_1 regularisation

The case of L_1 regularization is particularly interesting. Indeed, in this case, the derivative of $\mathcal{U}$ *w.r.t.* x_i is piecewise constant, thus $\mathcal{J}_i$ is a piecewise rational fraction of the third order for the numerator and the second for the denominator. An illustration of this piecewise fraction is given on Fig. 2 (for coefficients obtained from simulations). The derivative of $\mathcal{J}_i$ is a piecewise rational fraction of the fourth order, thus the local minima of $\mathcal{J}_i$ corresponds to the roots of fourth order polynomials. Its global minimum can be obtained explicitly with a low computational cost:

1. The whole $\mathbb{R}$ has to be divided in $n+1$ zones Z^k (for an order n neighborhood) depending on the positivity of $|x_i-x_j|$ for each j neighbor of i.
2. In each zone Z^k, the values x_i canceling the derivative of $\mathcal{J}_i$ are computed. They correspond to the roots of a fourth order polynomial and can be calculated explicitly. Only the roots which belong to the zone Z^k are selected.
3. The criterion is evaluated at the value of the selected roots as well as at the value of the neighbors of i. The value minimizing $\mathcal{J}_i$ corresponds to its global minimum.

4.2. L_2 regularisation

The L_2 regularization is usually easier to take into account than L_1 regularization. For the proposed algorithm, it is the opposite as the L_2 regularisation leads to a fifth degree rational fraction for the derivative of $\mathcal{J}_i$. Thus the computation of the minimum (root of a fifth degree polynomial) cannot be done explicitly and has to be computed numerically.

4.3. Huber potential function

The Huber function, as other convex potential functions, has an L_2 behavior near the origin and an L_1 behavior for large transition values, thus chosing such a model corresponds to a compromise between the L_1 and L_2 regularizations. This potential function can be used in the proposed algorithm as its derivative is piecewise polynomial (of degree zero or one). It leads to compute the root of a fifth degree piecewise polynomial defined over $3n+1$ zones for a size n neighborhood.

4.4. Other potential functions

The proposed algorithm is not restricted to potential functions whose derivatives are piecewise rational fractions; this property simply transforms the 1-D global optimization problem into a problem of polynomials roots calculation. Any other function can be taken into account but then requires implementation of an 1-D global optimization technique.

Such an algorithm remains interesting as the 1-D criterion is evaluated with a very low computational cost.

4.5. Additional constraints

Note that it is particularily easy to account for constraints on the values of the pixels in single site update algorithms while it is very difficult for global update algorithms. In our case, prior information such as $\boldsymbol{x}(i) \in [a_i, b_i]$ (or the positivity of the pixels), can be incorporated easily in the 1-D global optimisation technique.

5. SIMULATION RESULTS

The obtained solution at convergence corresponds to a global minimum of $\mathcal{J}_{\text{MAP}}$ with respect to each pixel, and to a local minimum of $\mathcal{J}_{\text{MAP}}$ if the potential fuction is differentiable. Nevertheless, there is no guarantee for this solution to correspond to the global minimum of $\mathcal{J}_{\text{MAP}}$ and the solution depends on the initialization and on the chosen sweep order of the pixels.[4] In spite of this property, the algorithm has given very satisfactory simulation results.

First, we compare the solution given by the proposed algorithm and by the conjugate gradient (CG) for an easy configuration. Both algorithms converge towards the same solution as the criterion seems to be unimodal, but the proposed SSUA converges faster than the CG. As an illustration, we represent in Fig. 3 the real part of the original object and of the solution given by the proposed and the CG algorithms after 10 iterations. As noticed and justified in [4] in a linear case, a local update strategy allows to recover faster the high frequencies of an object than a global update strategy (CG) which is noticeable in these results.

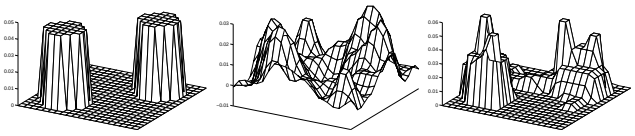

Figure 3: Original object (left) and solution given by the conjugate gradient (center) and by the proposed algorithm (left) after 10 iterations ($L_{1,1}$ regularization). The object has a relative contrast of 2, 21×21 pixels and 8×8 data are available with SNR = 20dB.

Then, we show some results in a more difficult configuration for which the CG get stuck in a local minimum. We represent in Fig. 4 the criterion value $\mathcal{J}_{\text{MAP}}(\boldsymbol{x}_n)$ and the relative reconstruction error ${}_2(\boldsymbol{x}_n) \triangleq \frac{\|\boldsymbol{x}-\boldsymbol{x}^*\|}{\|\boldsymbol{x}^*\|}$. One can notice that the algorithm seems to escape from a local minimum of the criterion $\mathcal{J}_{\text{MAP}}$ around the twentieth iteration.

6. CONCLUSION

We propose a deterministic single site update algorithm to compute a regularized solution to the nonlinear inverse problem of diffraction tomography.

[4] We have noticed that a checkerboard sweeping strategy is preferable to usual raster scanning, as its seems to avoid more local minima.

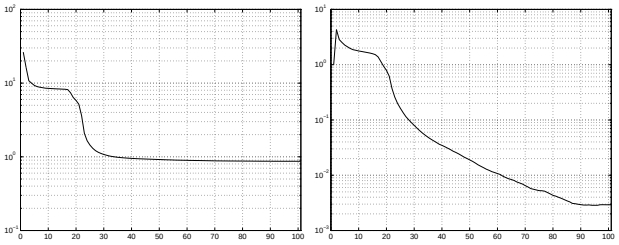

Figure 4: Value of the criterion (left), of the relative reconstruction error (center) and solution (right) given by the proposed SSUA. ($L_{1,1}$ regularisation, positivity constraint) The object has a relative contrast of 5.5, 11×11 pixels and 8×8 data are available with SNR = 20dB.

For each pixel, the 1-D global minimization can be performed at a very low computational cost, particularly for certain potential functions. Moreover, the updates between each minimization do not need to compute the forward problem which is computationally very expensive.

The algorithm has been shown to be more efficient than local optimization techniques such as the conjugate gradient, when the criterion does not have local minima. Moreover, it is able to reach the global minimum in more difficult configurations.

7. REFERENCES

[1] A. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*, IEEE Press, New York, NY, 1987.

[2] H. Carfantan and A. Mohammad-Djafari, "An overview of nonlinear diffraction tomography within the Bayesian estimation framework", in *Conference on Inverse Problems of Wave Propagation and Diffraction*, Aix-les-Bains, France, September 1996, INRIA.

[3] H. Carfantan and A. Mohammad-Djafari, "A Bayesian approach for nonlinear inverse scattering tomographic imaging", in *Proceedings of IEEE ICASSP*, Detroit, U.S.A., May 1995, vol. IV, pp. 2311–2314.

[4] K. Sauer and C. Bouman, "A local update strategy for iterative reconstruction from projections", *IEEE Transactions on Signal Processing*, vol. SP-41, no. 2, pp. 534–548, February 1993.

[5] J. E. Besag, "On the statistical analysis of dirty pictures (with discussion)", *Journal of the Royal Statistical Society B*, vol. 48, no. 3, pp. 259–302, 1986.

[6] S. Brette and J. Idier, "Optimized single site update's algorithms for image deblurring", Lausanne, September 1996, Internationnal Conference on Image Processing.

[7] H. Carfantan, *Approche bayésienne pour un problème inverse non linéaire en imagerie à ondes diffractées*, PhD thesis, Université de Paris-Sud Orsay, December 1996.