



APPROCHE BAYÉSIENNE ET ALGORITHME MULTIRÉSOLUTION POUR UN PROBLÈME INVERSE NON LINÉAIRE EN TOMOGRAPHIE DE DIFFRACTION

Hervé Carfantan et 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

RÉSUMÉ

Nous nous intéressons à un problème inverse non linéaire : la résolution des équations couplées en tomographie de diffraction. Nous avons récemment présenté la résolution de ce problème dans le cadre de l'estimation bayésienne [1]. Le calcul de l'estimateur du maximum *a posteriori* nécessite la minimisation d'un critère multimodal pour laquelle nous avons introduit une technique de relaxation déterministe s'inspirant du principe de non-convexité graduelle. Cette méthode donne des résultats intéressants comparée à d'autres méthodes existantes, elle est cependant très coûteuse en temps de calcul. Nous proposons ici de réduire le coût de cet algorithme de relaxation déterministe à l'aide d'un schéma multirésolution

1. INTRODUCTION

La reconstruction d'image en tomographie à ondes diffractées intervient dans de nombreux domaines tels que l'imagerie médicale, le contrôle non destructif et la géophysique. Elle vise à reconstruire des caractéristiques physiques d'un objet (la permittivité complexe) à partir de mesures d'un champ propagé à travers cet objet. La relation liant les mesures à l'objet recherché est non linéaire et a longtemps été linéarisée en utilisant les approximations de Born ou de Rytov dont la validité est limitée.

Dans un précédent travail [1] nous avons présenté la résolution de ce problème inverse non linéaire dans le cadre de l'estimation bayésienne, qui permet de combiner notre connaissance sur le bruit de mesure et sur l'objet à reconstruire. La solution, définie comme étant l'estimée au sens du *maximum a posteriori* (MAP), est une solution régularisée du problème. Le calcul de cet estimateur nécessite de minimiser un critère multimodal pour lequel les techniques d'optimisation globale de type recuit simulé ne peuvent être utilisées. Pour effectuer cette minimisation nous avons introduit une méthode de relaxation déterministe s'inspirant du principe de non-convexité graduelle, introduit par Blake et Zisserman [2] pour la segmentation d'image et généralisé aux problèmes inverses linéaires par Nikolova *et al.* [3]. Cet algorithme est sous optimal mais donne de bons résultats expérimentaux. Cependant il nécessite la minimisation locale de critères, dont le coût en

ABSTRACT

In order to solve the coupled equations in diffraction tomographic imaging, which is a non linear inverse problem, we presented recently a Bayesian reconstruction method [1]. As the determination of the MAP estimator requires the minimisation of a multimodal criterion, we introduced a deterministic relaxation scheme based on the Graduated Non Convexity principle to perform it. The method in [1] gave some good practical results compared to other existing methods, but with a heavy computational cost. In this paper, we propose to reduce this cost using a multiresolution scheme.

calcul est important.

Dans cette communication nous proposons une méthode qui consiste en (i) une modélisation multirésolution de notre problème, (ii) une modélisation *a priori* markovienne hiérarchique, et (iii) une relaxation déterministe aux différents niveaux de résolution. Le gain en temps de calcul est alors important par rapport à la même relaxation effectuée à un seul niveau de résolution.

2. MODÉLISATION DU PROBLÈME

La configuration géométrique du problème est présentée en Figure 1.a. Le but poursuivi est de reconstruire la permittivité complexe x d'un objet sur un domaine bidimensionnel D_o , à partir de mesures du champ diffracté y qui s'est propagé à travers cet objet, recueillies sur un ensemble discret de points D_m .

Les équations de Maxwell conduisent aux équations intégrales couplées :

$$\begin{cases} y(z_i) = \iint_{D_o} G_m(z_i, z') \phi(z') x(z') dz', & z_i \in D_m, \\ \phi(z) = \phi_0(z) + \iint_{D_o} G_o(z, z') \phi(z') x(z') dz', & z \in D_o, \end{cases}$$

dans lesquelles ϕ et ϕ_0 sont respectivement les champs total et incident sur le domaine D_o , et les fonctions G_m et G_o les fonctions de Green correspondant au domaine homogène dans lequel est plongé l'objet.



La discrétisation de ces équations par une méthode des moments nous permet de modéliser l'objet par un vecteur \mathbf{x} de dimension n correspondant à la projection de x sur une famille finie de fonctions $(\psi_k)_{k \in \{1 \dots n\}}$ formant un recouvrement du domaine D_o . Les équations, une fois discrétisées, peuvent s'écrire :

$$\begin{cases} \mathbf{y} &= \mathbf{G}_m \mathbf{X} \phi, \\ \phi &= \phi_0 + \mathbf{G}_o \mathbf{X} \phi, \end{cases}$$

où \mathbf{X} est une matrice carrée diagonale comportant les éléments de \mathbf{x} sur la diagonale, et les matrices \mathbf{G}_m , \mathbf{G}_o correspondent aux fonctions de Green pour la configuration choisie. \mathbf{y} est un vecteur de dimension m dans lequel sont stockées les mesures, et l'on a en général $n \gg m$, ce qui implique que ce système d'équations algébriques est largement sous-déterminé.

Le problème direct consiste à calculer les mesures \mathbf{y} correspondant à un objet \mathbf{x} . Ce problème est non linéaire du fait de la deuxième équation qui est implicite vis-à-vis de ϕ . On peut cependant écrire les deux équations couplées à l'aide d'une seule équation de la forme $\mathbf{y} = A(\mathbf{x})$, avec :

$$A(\mathbf{x}) = \mathbf{G}_m \mathbf{X} (\mathbf{I} - \mathbf{G}_o \mathbf{X})^{-1} \phi_0.$$

Le problème inverse consiste, quant à lui, à déterminer l'objet \mathbf{x} inconnu à partir des mesures bruitées \mathbf{y} . Ce problème est intrinsèquement mal posé et nécessite d'être régularisé pour donner une solution satisfaisante.

3. RÉOLUTION DU PROBLÈME

3.1. Définition d'un critère objectif

Nous avons proposé récemment [1] la résolution de ce problème dans le cadre de l'estimation bayésienne. Cette approche permet de combiner l'information contenue dans les mesures, par ailleurs insuffisantes à elles seules pour reconstruire l'objet, et l'information *a priori* sur cet objet.

Pour prendre en compte les incertitudes sur les mesures, on considère l'équation :

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

où le vecteur \mathbf{b} représentant les erreurs (de modélisation, de discrétisation ainsi que celles dues au système de mesure), est supposé additif, blanc, gaussien $\mathbf{b} \propto N(\mathbf{0}, \sigma^2 \mathbf{I})$, de variance connue σ^2 . De cette hypothèse on peut déduire la vraisemblance des mesures :

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

L'information *a priori* est introduite sous la forme d'une loi de probabilité :

$$p(\mathbf{x}) \propto \exp(-\mu U(\mathbf{x})).$$

Le choix de la fonction d'énergie U est essentiel. En traitement d'image, U est souvent l'énergie de champs de Markov, ce qui permet de modéliser la corrélation entre pixels voisins. La fonction d'énergie s'exprime alors comme la somme de fonctions potentiel V_c agissant sur un ensemble de points voisins c (appelé clique) :

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

Dans ce travail, nous avons utilisé des champs de Gauss Markov généralisés [4] dont la fonction potentiel est de la forme :

$$V_c(\mathbf{x}) = |\mathbf{x}_j - \mathbf{x}_i|^p, (\mathbf{x}_j, \mathbf{x}_i) \in c, 1 < p \leq 2.$$

Pour p proche de 2, ils pénalisent fortement les différences importantes entre les valeurs des pixels voisins et permettent de reconstruire des objets à faibles variations ($p = 2$: champs de Gauss Markov) ; en revanche, pour p proche de 1, ils pénalisent moins de telles différences et peuvent être utilisés pour reconstruire des objets possédant des discontinuités.

La solution du problème inverse est alors définie comme l'estimateur du maximum *a posteriori* :

$$\hat{\mathbf{x}}_{MAP} = \arg \max_{\mathbf{x}} p(\mathbf{x}|\mathbf{y}),$$

dont la détermination se ramène, par l'intermédiaire de la règle de Bayes :

$$p(\mathbf{x}|\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{x})p(\mathbf{x})}{p(\mathbf{y})},$$

à la minimisation d'un critère de la forme :

$$J(\mathbf{x}) = \|\mathbf{y} - A(\mathbf{x})\|^2 + \lambda U(\mathbf{x}),$$

où λ joue le rôle de paramètre de régularisation réglant le compromis entre la fidélité aux données et la confiance en l'*a priori*.

3.2. Minimisation d'un critère multimodal

Le problème se ramène donc à la minimisation du critère régularisé $J(\mathbf{x})$. Des critères de la même forme ont été largement étudiés dans le cas d'opérateurs linéaires pour lesquels la première partie du critère $\|\mathbf{y} - A\mathbf{x}\|^2$ est convexe. Deux cas sont alors à considérer : si la partie correspondant à l'*a priori* $U(\mathbf{x})$ est convexe, le critère $J(\mathbf{x})$ est alors convexe (et donc unimodal) et un simple algorithme de descente du type gradient conjugué suffit pour le minimiser. Pour des *a priori* plus élaborés (entraînant $U(\mathbf{x})$ non convexe), le critère $J(\mathbf{x})$ est multimodal. Deux types d'algorithmes ont alors été employés afin de l'optimiser :

- Le recuit simulé a été introduit en restauration d'image par Geman et Geman [5]. Il met à jour successivement

chaque pixel de l'image en fonction de sa probabilité *a posteriori* conditionnellement aux valeurs des autres pixels. Il n'est réellement efficace que lorsque cette distribution possède un support réduit, ce qui engendre un faible coût de calcul pour chaque mise à jour.

- La non-convexité graduelle (GNC), s'inspire des méthodes de continuation en mathématiques. Elle a été introduite en débruitage et en segmentation d'image par Blake et Zisserman [2] et généralisé pour les problèmes inverses linéaires par Nikolova *et al* [3]. Le principe de cet algorithme est très simple. Il consiste à approcher $J(\mathbf{x})$ par une suite de critères continûment dérivables $J_n(\mathbf{x})$ qui convergent vers $J(\mathbf{x})$ pour $n \rightarrow \infty$, dont le premier $J_0(\mathbf{x})$ est convexe. Alors, à chaque étape n , on calcule $\hat{\mathbf{x}}_n$, minimum local de J_n au voisinage de $\hat{\mathbf{x}}_{n-1}$. Pour tout point initial $\hat{\mathbf{x}}_0$, la suite $\hat{\mathbf{x}}_n$ converge vers un même minimum local de J qui est en pratique proche de son minimum global.

En ce qui nous concerne, nous nous retrouvons dans une situation renversée, à savoir que la fonction d'énergie $U(\mathbf{x})$ est convexe, mais le terme de fidélité aux données est multimodal du fait de la structure de l'opérateur A . De plus, le recuit simulé n'est pas applicable en pratique, en raison du large support de l'opérateur A et de son coût de calcul élevé. Dans [1] nous proposons une technique de relaxation déterministe sous optimale, inspirée du principe du GNC. Il s'agit dans un premier temps de construire une suite de critères $J_n(\mathbf{x})$ approchant $J(\mathbf{x})$. L'approximation de Born, longtemps utilisée pour résoudre ce problème, linéarise les équations couplées en négligeant le champ diffracté dans le domaine de l'objet ($\phi \approx \phi_0$); elle peut fournir une solution initiale pour notre schéma de relaxation. L'opérateur qui lui est associé :

$$A_0(\mathbf{x}) = \mathbf{G}_m \mathbf{X} \phi_0.$$

nous permet de débiter la relaxation par un critère convexe J_0 . Pour effectuer la relaxation de J_0 vers J , nous considérons la suite de critères :

$$J_n(\mathbf{x}) = \|\mathbf{y} - A_n(\mathbf{x})\|^2 + \lambda U(\mathbf{x}),$$

avec les opérateurs A_n définis par :

$$A_n(\mathbf{x}) = \mathbf{G}_m \mathbf{X} (I - r_n \mathbf{G}_o \mathbf{X})^{-1} \phi_0,$$

et r_n une suite de coefficients de relaxation telle que $r_0 = 0$ et $r_n \rightarrow 1, (n \rightarrow \infty)$. Cette suite de critères vérifie nos *desiderata*: J_0 est convexe et J_n tend vers J pour $n \rightarrow \infty$. En pratique, la suite de coefficients de relaxation r_n est prise de faible longueur (ex : $r_n = 0, 0.2 \dots, 1$).

Cette technique a fourni de bons résultats au regard d'autres méthodes existantes [1] mais reste cependant d'un coût de calcul élevé car elle requiert la minimisation locale de différents critères dont les coûts de calcul sont eux mêmes élevés.

4. MODÉLISATION ET ALGORITHME MULTIRÉSOLUTION

Le principe de tout schéma multirésolution «coarse to fine» est de considérer, dans un premier temps, le problème à une résolution grossière, puis de prendre en compte successivement des résolutions de plus en plus fines. De tels algorithmes ont montré leur intérêt dans de nombreux domaines et une synthèse de ces travaux dans le cadre de la modélisation markovienne a été réalisée dans [6]. Nous devons, afin d'utiliser un tel algorithme, modéliser le processus d'observation ainsi que l'informations *a priori* à différents niveaux de résolution.

4.1. Modélisation multirésolution du problème

La discrétisation des équations par la méthode des moments nous donne un vecteur \mathbf{x} correspondant à la projection de \mathbf{x} sur une famille finie de fonctions $(\psi_k)_{k \in \{1 \dots n\}}$. Nous n'allons pas nous intéresser ici au choix de ces fonctions (ψ_k) et nous les considérerons comme constantes sur un maillage rectangulaire du domaine de l'objet¹. En revanche, le nombre de ces fonctions est une donnée importante sur laquelle nous allons intervenir pour prendre en compte les observations à différents niveaux de résolutions. Remarquons que plus le vecteur \mathbf{x} est de grande dimension, plus l'évaluation du problème direct est coûteux en temps de calcul, ce qui souligne l'intérêt d'un algorithme multirésolution.

La discrétisation des équations à une résolution s donnée s'écrit :

$$\begin{cases} \mathbf{y} &= \mathbf{G}_m^{(s)} \mathbf{X}^{(s)} \phi^{(s)}, \\ \phi^{(s)} &= \phi_0^{(s)} + \mathbf{G}_o^{(s)} \mathbf{X}^{(s)} \phi^{(s)}. \end{cases}$$

Les matrices $\mathbf{G}_m^{(s)}$ et $\mathbf{G}_o^{(s)}$ correspondant aux fonctions de Green à la résolution s sont calculées une fois pour toute pour une configuration donnée. La correspondance entre l'objet à une résolution s et l'objet à une résolution plus grossière $s - 1$ pourra être effectuée par une simple décimation par 2. Cela nous amène donc à considérer une pyramide multirésolution (Figure 1.b) de discrétisation de l'objet.

4.2. Modélisation Markovienne hiérarchique de l'*a priori*

Nous utilisons un modèle markovien hiérarchique [6] afin de modéliser l'information *a priori* à différents niveaux de résolutions. De l'expression des champs de Gauss Markov

1. Un autre choix pour ces fonctions, une famille d'ondelettes par exemple, peut engendrer des difficultés lors du calcul des matrices \mathbf{G}_m et \mathbf{G}_o , mais peut surtout créer des problèmes pour l'interprétation de la régularisation. En effet, l'information *a priori* est introduite sur les coefficients de la projection de \mathbf{x} sur la famille $(\psi_k)_{k \in \{1 \dots n\}}$, et il n'est pas toujours facile de modéliser cette information pour toute famille de fonctions.



généralisés présentés précédemment, on peut déduire leur fonction d'énergie à une résolution donnée :

$$U(\mathbf{x}^{(s)}) = \sum_{i,j \in V_i^{(s)}} |\mathbf{x}_j^{(s)} - \mathbf{x}_i^{(s)}|^p,$$

pour des voisinages $V_i^{(s)}$ et $1 < p \leq 2$. Il n'est pas nécessaire dans notre situation de considérer un modèle hiérarchique plus complexe, en particulier de lien entre les différentes résolutions, comme c'est souvent le cas pour des problèmes de segmentation ou de restauration d'image. En effet, dans ces situations, les données sont présentes sous la même forme que l'image, éventuellement bruitées et lissées ; un lien entre le modèle aux différentes résolutions et celui à la résolution de l'image est donc nécessaire voire indispensable [6]. Mais dans le cadre de la reconstruction d'image, les données peuvent être utilisées telles quelles pour tous les niveaux de résolution et l'on peut, en pratique, vouloir reconstruire l'objet avec un nombre quelconque de pixels.

4.3. Algorithme de relaxation multirésolution

Ayant explicité le modèle d'observation et le modèle *a priori* aux différents niveaux de résolution, il nous reste à préciser l'algorithme de relaxation multirésolution utilisé. Nous effectuons, dans un premier temps, la relaxation du critère J à une résolution grossière s pour laquelle le coût de calcul des critères $J_n^{(s)}$ est peu élevé. Puis nous propageons les résultats de cette relaxation à une résolution plus fine $s + 1$, ce qui nous donne une solution initiale pour la minimisation locale du critère $J^{(s+1)}$. Enfin nous itérons ce processus jusqu'à atteindre la résolution la plus fine désirée.

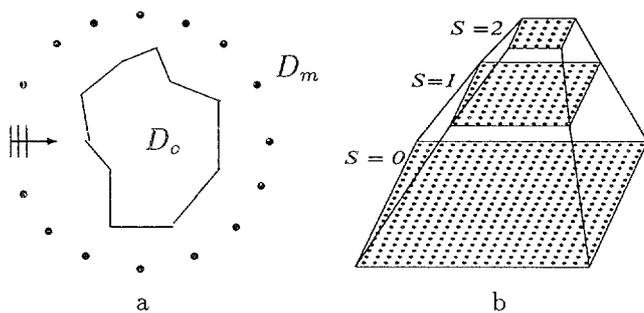


Figure 1:

a) Configuration, b) pyramide multirésolution.

5. SIMULATIONS ET CONCLUSIONS

La Figure 2 présente un exemple de résultats de simulation de cet algorithme de relaxation déterministe multirésolution. L'objet est de taille 21×21 pixels, il est modélisé par un champ de Gauss Markov Généralisé avec $p = 1$, 1 et des voisinages verticaux et horizontaux d'ordre 1. Les mesures sont au nombre de $8 \times 8 = 64$ et sont bruitées avec un rapport signal à bruit de 20dB. Nous ne montrons ici que la partie réelle de l'objet.

La relaxation est effectuée en six étapes à une résolution deux fois plus grossière que la résolution voulue, puis les résultats sont propagés jusqu'à la résolution la plus fine.

Le résultat ainsi obtenu est le même que lorsque la relaxation est effectuée toujours à la même résolution mais avec un temps de calcul 4 fois moins élevé.

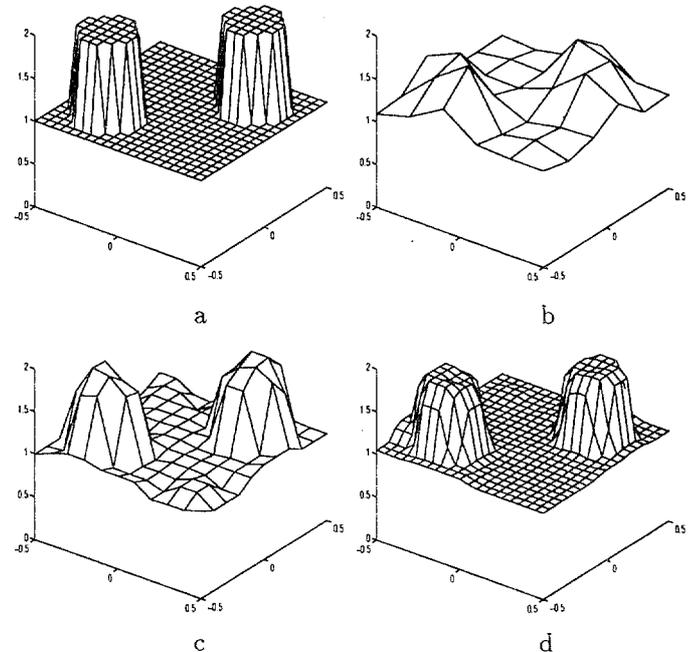


Figure 2:

Reconstructions : a) Objet original, b) Résolution grossière, c) Résolution intermédiaire, d) Résolution finale.

6. RÉFÉRENCES

- [1] H. Carfantan, et A. Mohammad-Djafari, "A Bayesian Approach for Nonlinear Inverse Scattering Tomographic Imaging," *IEEE ICASSP*, vol. 4, p. 2311-2314, (1995).
- [2] A. Blake, et A. Zisserman, "Visual Reconstruction," *MIT Press, Cambridge MA*, (1987).
- [3] M. Nikolova, A. Mohammad-Djafari, et J. Idier, "Inversion of large-support ill-conditioned linear operators using a Markov model with a line process," *IEEE ICASSP*, vol. 5, p. 357-360, (1994).
- [4] Ch. Bouman, et K. Sauer, "A generalized Gaussian image model for edge-preserving MAP estimation," *IEEE MI*, vol. 2, p. 296-310, (1993).
- [5] S. Geman, et D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 6, p. 721-741, (1984).
- [6] C. Graffigne, F. Heitz, F. Prêteux, M. Sigelle, et J. Zerubia "Modèles markoviens hiérarchiques pour l'analyse d'images," *Rapport du GDR TdSI*, Novembre 1994.