

ÉCOLE NATIONALE SUPÉRIEURE
DES
TÉLÉCOMMUNICATIONS DE BRETAGNE
Option Signal et Image
Promotion 2002

UNIVERSITÉ DE RENNES I
DEA Signal Télécommunications Image Radar
Option Image

PROJET DE FIN D'ÉTUDES
AU
LABORATOIRE DES SIGNAUX ET SYSTÈMES
(CNRS-SUPELEC-UPS)
Mai-Septembre 2002

FUSION BAYÉSIENNE DE DONNÉES EN
RECONSTRUCTION TOMOGRAPHIQUE
APPLICATION DES MÉTHODES DE MONTE-CARLO

Christophe LAPLANCHE

Responsable de stage : Ali MOHAMMAD-DJAFARI
Tuteur de stage : Jean-José JACQ
Responsable d'option : Guy CAZUGUEL
Responsable de DEA : Jean-Marc BOUCHER

Remerciements

J'envoie tout d'abord un grand merci à Ali, mon responsable de stage, pour avoir accompagné mes quelques premiers pas dans le monde de la recherche. Merci à toute sa disponibilité, sa patience, et sa gentillesse surtout.

Merci aux deux thésards du Groupe des Problèmes Inverses, Marc et Hichem, qui ont toujours répondu avec soin à la moindre de mes questions.

Merci à Mehdi, autant pour ses tapes dans le dos de bon matin, que pour tous ses coups de pouces. C'est la plus enrichissante des thèses au sein du GPI que je lui souhaite!

Puis...

A maman, pour tous tes mots justes.

Et Armony... je ne te remercie pas pour ton amour... ou je vais prendre une bonne gifle. Ces premiers mois dans le monde scientifique ont aussi été les premiers passés à tes côtés. Les plus diff... heureux! La soutenance du 17 septembre sera la veille de ton anniversaire, et permets-moi de t'offrir ce petit compte-rendu...

Introduction

Nous vivons dans un monde de radioactivité, qui nous soumet en permanence à une irradiation naturelle, estimée, en France, de 2 à 3 milliSievert (mSv) par an en moyenne. Si l'on ajoute la radioactivité artificielle (42%), essentiellement d'origine médicale (41%), l'irradiation moyenne atteint 4 mSv par an, pour chaque individu.

A l'heure actuelle, les effets sur la santé humaine d'une exposition à des doses inférieures à 100mSv font l'objet de débats scientifiques¹. Néanmoins, la limite de dose, pour le personnel professionnellement exposé, est de 100 mSv *pour 5 ans*, soit 20 mSv par an, avec un maximum toléré de 50 mSv pour une année donnée.

Les scanners irradient beaucoup plus les patients que la radiologie conventionnelle². Un scanner céphalique, par exemple, administre une dose de 40 mSv. Le radiodiagnostic se voit ainsi refusé à certains patients qui l'exigent.

Dans les scanners médicaux, les méthodes utilisées en reconstruction exigent un nombre important de mesures (soit autant de faisceaux de photons X traversant le patient) qui peut être diminué par l'emploi d'un autre algorithme de reconstruction, et par l'introduction de données obtenues par d'autres systèmes de mesure (non ionisants). C'est ce que permet la fusion de données dans le cadre d'une estimation bayésienne, exposée dans ce compte-rendu de stage.

Y sera tout d'abord présenté le principe de la tomographie à rayons X, et son opérateur de projection, la *transformée de Radon* (§1.1.1). Son inversion théorique a donné naissance à de nombreuses méthodes de reconstruction (§1.2), où apparaîtra l'important opérateur de *retroprojection* (§1.4), très souvent utilisé dans les méthodes d'estimations présentées. La reconstruction se fera par utilisation de divers estimateurs bayésiens (§1.5).

Le modèle présenté (§2.2.1) est une réécriture homogène de sa version inhomogène [1]. Quelques petites précisions théoriques (annexes) permettront alors l'écriture des lois (§2.2.2) et des algorithmes (§2.2.3 et §2.2.4) nécessaires à l'estimation.

Le bruit additionnel sera étudié, par étude du rôle dans les algorithmes, puis l'estimation, des variances le caractérisant, ceci par application au problème des *méthodes de Monte-Carlo* (§2.2.4). Ces estimations permettront une reconstruction stable, et une évaluation du rapport signal sur bruit des données tomographiques.

Enfin, l'importance de l'introduction des données échographiques dans l'aide à l'estimation sera mise en évidence (§2.3.2).

¹s'il existe une relation linéaire *vérifiée* entre exposition aux rayonnements ionisants et excès de cancers (leucémies et cancers solides), celle-ci n'a pas été démontrée, à ce jour, pour les très faibles doses. Mais on ne peut actuellement exclure un excès de leucémies par exposition *in utero* de l'ordre de 10 à 20 mSv.

²d'un facteur 20 pour un examen du thorax, du cliché standard à la tomodensitométrie.

Table des matières

| | |
|---|-----------|
| Remerciements | 3 |
| Introduction | 4 |
| 1 Reconstruction tomographique | 6 |
| 1.1 Introduction à la tomographie X | 6 |
| 1.1.1 De la constante d'atténuation à la Transformée de Radon | 6 |
| 1.1.2 Le tomographe à rayons X | 7 |
| 1.2 Méthodes analytiques de reconstruction | 8 |
| 1.2.1 Inversion directe | 8 |
| 1.2.2 Convolution | 9 |
| 1.2.3 Synthèse de Fourier | 10 |
| 1.2.4 Limitations | 10 |
| 1.3 Discrétisation | 10 |
| 1.3.1 Méthodes algébriques | 10 |
| 1.3.2 Matrice des projections | 11 |
| 1.4 Rétroprojection | 11 |
| 1.4.1 L'opérateur continu de rétroprojection | 12 |
| 1.4.2 L'opérateur discret de rétroprojection | 12 |
| 1.4.3 Approche de l'inversion par rétroprojection | 12 |
| 1.5 Approche bayésienne en reconstruction | 13 |
| 1.5.1 Modèle | 13 |
| 1.5.2 Estimateur bayésien | 13 |
| 1.5.3 Loi a posteriori | 14 |
| 1.5.4 Estimation des inconnues | 14 |
| 1.6 Influence du bruit | 15 |
| 2 Fusion homogène en reconstruction tomographique | 16 |
| 2.1 Reconstruction à partir d'un faible nombre de projections | 16 |
| 2.2 Modèle et méthodes d'estimation | 16 |
| 2.2.1 Modèle | 17 |
| 2.2.2 Ecriture des lois <i>a posteriori</i> | 19 |
| 2.2.3 Estimation par MAP | 20 |
| 2.2.4 Estimation par méthode de Monte-Carlo | 22 |
| 2.3 Simulations | 25 |
| 2.3.1 MAP VS MCMC | 25 |
| 2.3.2 Algorithmes hybrides | 29 |

| | |
|---|-----------|
| Conclusion | 32 |
| Annexes | 32 |
| A Champs de Gibbs-Markov | 33 |
| B Gradients matriciels | 36 |
| C Estimation de la variance d'une loi gaussienne | 39 |
| D Méthodes de Monte-Carlo | 41 |
| E Symétrie des opérateurs de dérivation | 43 |
| F JMAP hiérarchique | 46 |
| Bibliographie | 47 |

Chapitre 1

Reconstruction tomographique

1.1 Introduction à la tomographie X

1.1.1 De la constante d'atténuation à la Transformée de Radon

Considérons un objet de densité uniforme, traversé par un photon d'énergie ϵ connue. On définit la *constante d'atténuation linéaire* de cet objet par :

$$\mu_\epsilon = -\frac{1}{l} \ln(\rho) \quad (1.1)$$

où : l est la longueur de matière traversée par le photon,
 ρ la probabilité que le photon ne soit ni absorbé, ni dévié par la matière constituant l'objet (FIG.1.1).

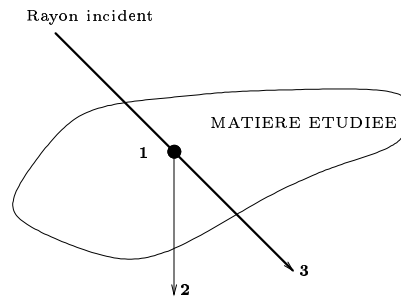


FIG. 1.1 – Photon absorbé (1), dévié (2) ou ni absorbé ni dévié (3)

Cette grandeur dépend fortement de l'énergie ϵ du photon (la probabilité qu'un photon traverse sans encombre une feuille de papier est faible dans le cas du visible, forte dans le cas micro-onde), et s'approche (§2.2) de la *densité de la matière* dans le cas X.

Inversement, considérons un objet inhomogène de constante d'atténuation μ_ϵ connue en son sein. La probabilité qu'un photon d'énergie ϵ le traverse sans être dévié ni absorbé sur le chemin L est :

$$\rho = \exp\left(-\int_L \mu_\epsilon(L) dL\right) \quad (1.2)$$

Ainsi, pour un faisceau mono-énergétique, étroit, de direction L , d'énergie ϵ , d'intensité $I_0(L)$ en entrée et $I(L)$ en sortie de l'objet, nous obtenons la loi de *Beer-Lambert* :

$$I(L) = I_0(L) \exp \left(- \int_L \mu_\epsilon(L) dL \right) \quad (1.3)$$

En se limitant désormais à un plan de hauteur fixée $P : z = z_0$, la droite L est alors entièrement définie par sa distance r à l'origine et son angle Φ par rapport à l'axe Ox (FIG.1.2). La *projection*

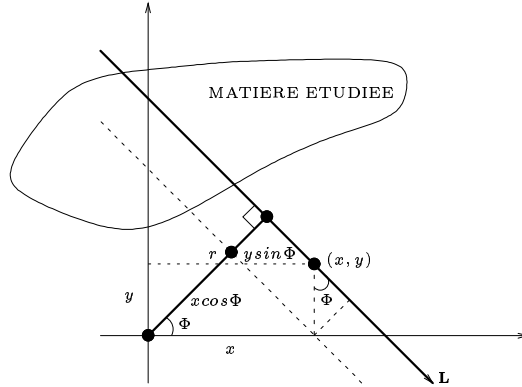


FIG. 1.2 – Direction de projection $L : r = x \cos \phi + y \sin \phi$

de l'objet dans la direction (r, Φ) est définie comme la quantité mesurée :

$$p(r, \Phi) = - \log \frac{I(L)}{I_0(L)} = \int_L \mu_\epsilon(L) dL \quad (1.4)$$

L'équation cartésienne de $L : r = x \cos \phi + y \sin \phi$ permet alors d'écrire :

$$p(r, \Phi) = \iint \mu_\epsilon(x, y) \delta(r - x \cos \phi - y \sin \phi) dx dy \quad (1.5)$$

La fonction $p(r, \Phi)$ est la *transformée de Radon* de la fonction $\mu_\epsilon(x, y)$. La mesure de $p(r, \Phi)$ pour la direction L fournit donc un échantillon de la transformée de Radon de la fonction atténuation μ_ϵ de la région observée.

1.1.2 Le tomographe à rayons X

La *reconstruction* d'image en tomographie X consiste à rechercher une estimation de $\mu_\epsilon(x, y)$ à partir des projection $p(r, \Phi)$. Le tomographe à rayons X effectue une succession de mesures de ces valeurs $p(r, \Phi)$ (FIG.1.3). Une source X effectue une rotation autour de l'objet à étudier. A chacune de ses positions correspond un ensemble de mesures de projections (une *vue*). Ces *vues* sont alors regroupées et représentées dans un *sinogramme* (FIG.2.2). La tomographie X est un cas particulier de rendu tomographique, lui-même procédé spécifique de *reconstruction à partir de projections*, qui reste un *problème d'inversion*. Ces méthodes sont utilisées dans des domaines aussi variés que la radio-astronomie, la géophysique, le contrôle non destructif (médecine, avionique, sylviculture, ...), ...

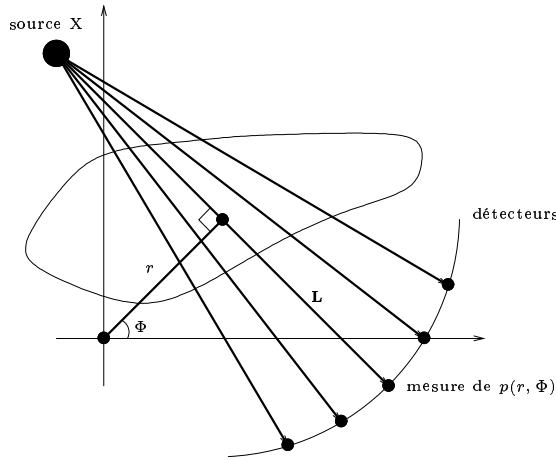


FIG. 1.3 – Mesures des projections dans le cas d'un tomographe à rayons X

1.2 Méthodes analytiques de reconstruction

Considérons une fonction en coordonnées cartésiennes $f(x, y) : \mathbb{R}^2 \mapsto \mathbb{R}$. La *transformée de Radon* de f est définie comme la fonction en coordonnées polaires $p(r, \Phi)$:

$$p(r, \Phi) = \iint f(x, y) \delta(r - x \cos \phi - y \sin \phi) dx dy \quad (1.6)$$

et son inverse est donnée par :

$$f(x, y) = \frac{1}{2\pi^2} \int_0^\pi \int_0^\infty \frac{\partial p(r, \Phi) / \partial r}{r - x \cos \phi - y \sin \phi} dr d\Phi \quad (1.7)$$

La transformée de Radon d'une fonction $f : \mathbb{R}^n \mapsto \mathbb{R}$ peut aussi être définie comme somme de projections sur des hyperplans, et son inverse calculée analytiquement [6]. Seul le cas bidimensionnel sera étudié par la suite.

Les *méthodes analytiques de reconstruction* consistent à exprimer une estimée de f à partir d'une *interprétation* en terme de succession d'opérateurs de la transformée de Radon inverse.

1.2.1 Inversion directe

La transformée de Radon inverse de $p(r, \Phi)$ peut se mettre sous la forme :

$$f(x, y) = -\frac{1}{2\pi} \mathcal{B} \cdot \mathcal{H} \cdot \mathcal{D}[p(r, \Phi)] \quad (1.8)$$

où : \mathcal{D} est l'*opérateur de dérivation* par rapport à la première variable (r),
 \mathcal{H} est la *transformée de Hilbert* par rapport à la première variable (r),
 \mathcal{B} est l'*opérateur de rétroprojection* (§1.4) :

$$\mathcal{H}[g(t)] = -\frac{1}{\pi} \int_{\mathbb{R}} \frac{g(t')}{t' - t} dt' \quad (1.9)$$

$$\boxed{\mathcal{B}[g(r, \Phi)](x, y) = \int_0^\pi g(x \cos \Phi + y \sin \Phi, \Phi) d\Phi} \quad (1.10)$$

La méthode d'inversion directe consiste à appliquer ces trois opérateurs sur la fonction de projection $p(r, \Phi)$:

$$\boxed{p(r, \Phi) \longmapsto \boxed{\mathcal{D}} \longmapsto \boxed{\mathcal{H}} \longmapsto \boxed{\mathcal{B}} \longmapsto f(x, y)}$$

1.2.2 Convolution

La transformée de Radon inverse peut s'écrire sous forme de *rétroprojection d'un filtrage* ou d'une *rétroprojection filtrée*

La rétroprojection s'approche donc (à un filtrage près) d'une inversion directe (§1.4.3).

rétroprojection d'un filtrage

$$\boxed{f(x, y) = \mathcal{B} \cdot \mathcal{C}_1[p(r, \Phi)]} \quad (1.11)$$

où $\mathcal{C}_1 = \mathcal{H} \cdot \mathcal{D}$ est le filtrage 1D de réponse fréquentielle $|\Omega|$:

$$\mathcal{C}_1[p(r, \Phi)] = (h_1 \star_r p)(r, \Phi) \quad \text{avec} \quad \mathcal{F}_1[h_1(r)](\Omega) = |\Omega| \quad (1.12)$$

pour \mathcal{F}_1 la *transformée de Fourier* par rapport à la première variable (r).

Cette méthode consiste donc à appliquer à $p(r, \Phi)$ un filtre passe-haut de réponse fréquentielle $|\Omega|$ avant de rétroprojeter :

$$\boxed{p(r, \Phi) \longmapsto \boxed{\mathcal{F}_1} \longmapsto \boxed{|\Omega|} \longmapsto \boxed{\mathcal{F}_1^{-1}} \longmapsto \boxed{\mathcal{B}} \longmapsto f(x, y)}$$

rétroprojection filtrée

$$\boxed{f(x, y) = \mathcal{C}_2 \cdot \mathcal{B}[p(r, \Phi)]} \quad (1.13)$$

où \mathcal{C}_2 est le filtrage 2D de réponse fréquentielle $|\Omega| = \sqrt{\omega_x^2 + \omega_y^2}$:

$$\mathcal{C}_2[g(x, y)] = (h_2 \star g)(x, y) \quad \text{avec} \quad \mathcal{F}_2[h_2(x, y)](\omega_x, \omega_y) = \sqrt{\omega_x^2 + \omega_y^2} \quad (1.14)$$

pour \mathcal{F}_2 la *transformée de Fourier* 2D.

Cette seconde méthode de convolution consiste donc à appliquer à la rétroprojection de $p(r, \Phi)$ un filtre passe-haut de réponse fréquentielle $|\Omega|$.

$$\boxed{p(r, \Phi) \longmapsto \boxed{\mathcal{B}} \longmapsto \boxed{\mathcal{F}_2} \longmapsto \boxed{|\Omega| = \sqrt{\omega_x^2 + \omega_y^2}} \longmapsto \boxed{\mathcal{F}_2^{-1}} \longmapsto f(x, y)}$$

1.2.3 Synthèse de Fourier

Le théorème de coupe-projection énonce que :

La transformée de Fourier (par rapport à la première variable r) de la transformée de Radon est la transformée de Fourier de l'image à reconstruire

Soit, avec les notations vues plus haut :

$$\mathcal{F}_1[p(r, \Phi)](\Omega, \Phi) = \mathcal{F}_2[f(x, y)](\omega_x, \omega_y) \quad (1.15)$$

avec :

$$\begin{cases} \omega_x = \Omega \cos \Phi \\ \omega_y = \Omega \sin \Phi \end{cases} \quad (1.16)$$

Cette dernière méthode analytique se réduit donc à :

$$p(r, \Phi) \longmapsto \boxed{\mathcal{F}_1} \longmapsto \boxed{\text{Interpolation } (\Omega, \Phi) \rightarrow (\omega_x, \omega_y)} \longmapsto \boxed{\mathcal{F}_2^{-1}} \longmapsto f(x, y)$$

1.2.4 Limitations

Toutes ces méthodes supposent $p(r, \Phi)$ parfaitement connu. Mais en pratique seul un nombre limité de projections (voire très limité en contrôle non destructif industriel) est accessible. Ces méthodes peuvent être approximées et discrétisées. Cependant [2] :

Une image ne peut être parfaitement reconstruite à partir d'un nombre limité de ses vues

Une infinité d'objets (plus ou moins éloignés) peuvent correspondre à un même ensemble de projections. Il devient alors nécessaire d'introduire de l'information supplémentaire pour aider à l'estimation. La modélisation discrète apparaît mieux adaptée à l'introduction de ces données *a priori*.

1.3 Discrétisation

1.3.1 Méthodes algébriques

Elles consistent à estimer un nombre fini de paramètres caractérisant *partiellement* l'image à reconstruire. L'objet peut être représenté par un *modèle paramétrique* (superposition d'ellipsoïdes dans le cas de la tomographie X en médical avec estimation des paramètres les caractérisant, ...), ou être *projeté sur une base de fonctions appropriées* (ondelettes, fonctions indicatrices, ...).

C'est la projection sur des fonctions indicatrices représentées par des pixels qui est utilisée ici

1.3.2 Matrice des projections

La projection tomographique devient une simple opération matricielle

En effet :

$$p(r_i, \Phi_j) = \sum_{(k,l)} f(k, l) \cdot h_{i,j}(k, l) \quad (1.17)$$

En notant : $f(k, l)$ la valeur moyenne de $f(x, y)$ sur le pixel (k, l) ,
 $h_{i,j}(k, l)$ la distance parcourue par le rayon (r_i, Φ_j) à travers le pixel (k, l) :

$$h_{i,j}(k, l) = \iint_{\text{pixel}(k,l)} \delta(r_i - x \cos \Phi_j - y \sin \Phi_j) dx dy \quad (1.18)$$

Soit, en réindiquant les pixels $(k, l) \in \llbracket 1, K \rrbracket \times \llbracket 1, L \rrbracket$ colonne par colonne par $n = (l - 1) \cdot K + k$,
 $N = K \cdot L$, en notant $f = (f_n)_{n \in \llbracket 1, N \rrbracket}$ et $h_{i,j} = (h_{i,j}(n))_{n \in \llbracket 1, N \rrbracket}$:

$$p(r_i, \Phi_j) = h_{i,j}^\top \cdot f \quad (1.19)$$

Puis en réindiquant les projections $(i, j) \in \llbracket 1, I \rrbracket \times \llbracket 1, J \rrbracket$ par $m = (j - 1) \cdot I + i$, $M = I \cdot J$, en
notant $p = (p(m))_{m \in \llbracket 1, M \rrbracket}$ et la *matrice des projections* $H = (h_m)_{m \in \llbracket 1, M \rrbracket}$:

$$\boxed{p = H \cdot f} \quad (1.20)$$

Ainsi : $N = K \cdot L$ est la taille de l'image originale, de K lignes et L colonnes,
 $M = I \cdot J$ est le nombre de projections, pour I directions angulaires et J détecteurs,
 f est l'image originale sous forme vectorielle,
 p est le vecteur des projections de f (son *sinogramme*, §1.1.2),
 H est la *matrice des projections*, de dimensions $M \cdot N$, indiquant la distance parcourue
par chacun des rayons (r_i, Φ_j) pour chacun des pixels (k, l) (§1.4.2 pour une
interprétation plus détaillée de H).

Pour une image de taille 256×256 , 128 positions angulaires et 64 détecteurs, H est de taille
536870912! En pratique, dans les scanners commerciaux du domaine médical, les images à re-
construire sont de taille 1024×1024 , pour 1024 positions angulaires et 1024 détecteurs, soit une
matrice de projections de l'ordre du Téra-octet. H n'est donc pas nécessairement carrée, et, même
dans ce cas, *mal conditionnée*, pour être inversée directement [5]. Mais inverser H reviendrait à
inverser la transformée de Radon pour un nombre dénombrable d'échantillons ce qui est, comme
celà a été vu au §1.2.4, impossible.

1.4 Rétroprojection

La procédure d'inversion de la transformée de Radon
peut être approchée par une simple rétroprojection

1.4.1 L'opérateur continu de rétroprojection

La droite $L(r, \Phi)$ a été définie plus haut (FIG.1.2) comme la droite distante de r à l'origine, et d'angle Φ par rapport à l'axe Ox . $L(x \cos \Phi + y \sin \Phi, \Phi)$ est donc la droite passant par (x, y) , d'angle Φ par rapport à l'axe Ox . L'opérateur de rétroprojection défini en (1.10) comme :

$$\mathcal{B}[g(r, \Phi)](x, y) = \int_0^\pi g(x \cos \Phi + y \sin \Phi, \Phi) d\Phi \quad (1.21)$$

fait donc la somme des projections (filtrées, §1.2.2) $g(r, \Phi)$ pour toutes les directions L passant par (x, y) .

1.4.2 L'opérateur discret de rétroprojection

En approche discrète de la reconstruction, la projection définie par le rayon (i, j) s'écrit (1.17) :

$$p(i, j) = \sum_{(k, l)} f(k, l) \cdot h_{i,j}(k, l) \quad (1.22)$$

Les éléments non nuls de la ligne $h_{i,j}$ de H donne donc l'ensemble (pondéré) des pixels (k, l) traversés par le rayon (i, j) . Inversement, les éléments non nuls d'une colonne (k, l) de H donne l'ensemble (pondéré) des rayons (i, j) traversant le pixel (k, l) .

$$\hat{f}(k, l) = \sum_{(i,j)} h_{i,j}(k, l) \cdot p(i, j) \quad \text{soit} \quad \hat{f} = H^T \cdot p \quad (1.23)$$

réalise donc l'opération de rétroprojection définie plus haut. Ainsi, en approche discrète :

H^T est l'opérateur de rétroprojection.

1.4.3 Approche de l'inversion par rétroprojection

La rétroprojection s'identifie comme l'opérateur adjoint de la transformée de Radon [6], soit :

$$\langle \mathcal{R}[f], p \rangle = \langle f, \mathcal{B}[p] \rangle \quad (1.24)$$

qui se vérifie très simplement dans le cas discret :

$$\begin{aligned} \langle Hf, p \rangle &= (Hf)^T p \\ &= f^T (H^T p) \\ &= \langle f, H^T p \rangle \end{aligned} \quad (1.25)$$

Cet opérateur adjoint est aussi, à un filtrage passe-haut près, la transformée de Radon inverse (§1.2.2).

La rétroprojection apparaît ainsi comme une version floue d'une inversion rigoureuse, ce qui justifie son utilisation en pratique (scanners du domaine médical) comme simple opérateur de reconstruction (FIG.1.4).

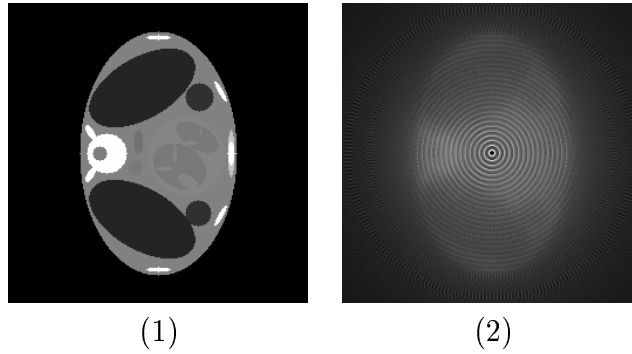


FIG. 1.4 – Fantome (1), reconstruction par rétroprojection (2)

1.5 Approche bayésienne en reconstruction

1.5.1 Modèle

Soit un champ mesuré y obtenu à partir d'un champ x , x et y reliés par le modèle :

$$\boxed{y = f_{\theta}(x) + \epsilon(\theta)} \quad (1.26)$$

où : θ est un ensemble de paramètres (supposés inconnus et à estimer) caractérisant le modèle,
 ϵ un bruit additif de loi connue.

L'estimation bayésienne consiste à estimer $z = (x, \theta)$ à partir des données y et d'*informations a priori*, introduites dans la *nature* des lois (bruit gaussien) des champs à estimer. Les *hyperparamètres* les définissant peuvent être fixés (moyenne nulle), estimés (variance qui suit une loi inverse-gamma de paramètres fixés).

1.5.2 Estimateur bayésien

La théorie de la décision assigne un coût $C(z, d)$ (z élément de l'espace des paramètres, d de celui des décisions) à chaque estimation. On peut définir le *coût moyen a posteriori* :

$$C(d|y) = E_{z|y}(C(z, d)) = \int C(\omega, d)p(\omega|y)d\omega \quad (1.27)$$

L'*estimateur bayésien* est alors celui qui minimise le coût moyen :

$$\hat{z} = \arg \min_d \{C(d|y)\} \quad (1.28)$$

Pour un *coût tout ou rien* $C(z, d) = 1 - \delta(z - d)$, l'estimateur de Bayes associé est l'argument qui *maximise la loi a posteriori* (MAP) :

$$\boxed{\hat{z} = \arg \max_z p(z|y)} \quad (1.29)$$

Pour un *coût linéaire* $C(z, d) = |z - d|$, l'estimateur de Bayes associé est la *médiane de la loi a posteriori*.

Pour un *coût quadratique* $C(z, d) = (z - d)^2$, l'estimateur de Bayes associé est la *moyenne a posteriori* (MMSE) :

$$\boxed{\hat{z} = E_{z|y}(z)} \quad (1.30)$$

C'est l'estimateur au sens du MAP qui sera étudié par la suite (et l'estimateur par MMSE sera utilisé implicitement dans les méthodes de Monte-Carlo, annexe D).

1.5.3 Loi a posteriori

Dans tous les cas, la loi *a posteriori* cherche à être exprimée. En utilisant la règle de Bayes :

$$p(z|y) = \frac{p(y|z).p(z)}{p(y)} \quad (1.31)$$

Soit (le champ y est connu) :

$$\boxed{p(z|y) \propto p(y|z).p(z)} \quad (1.32)$$

où : $p(y|z)$ est la *vraisemblance*, ici *loi du bruit* ϵ supposée connue (si $\epsilon \sim \mathcal{N}(0, \sigma^2)$,
 $y|z \sim \mathcal{N}(f_\theta(x), \sigma^2)$)
 $p(z)$ est la loi *a priori* sur les inconnues.

Si aucune information ne peut être apportée sur la loi *a priori* $p(z)$, une *loi uniforme* est choisie, et le MAP devient *maximum de vraisemblance* (MV). Dans le cas contraire, la loi *a posteriori* peut être exprimée (à une constante multiplicative près), et maximisée.

1.5.4 Estimation des inconnues

Inconnues x et paramètres θ peuvent être estimés simultanément, dans le cas de la *maximisation jointe* (JMAP) :

$$\hat{z} = \widehat{(x, \theta)} = \arg \max_{(x, \theta)} p(x, \theta|y) \quad (1.33)$$

Le champ x (les pixels d'une image à reconstruire) et les paramètres θ (du modèle et des lois *a priori*) sont de nature très différentes. La maximisation jointe peut s'avérer difficile. Inconnues et paramètres peuvent alors être estimés successivement, avec estimation de θ , puis x :

$$\begin{cases} \hat{\theta} = \arg \max_{\theta} p(\theta|y) \\ \hat{x} = \arg \max_x p(x|y, \hat{\theta}) \end{cases} \quad \text{avec } p(\theta|y) = \int_x p(x, \theta|y) dx \quad (1.34)$$

ou estimation de x , puis θ :

$$\begin{cases} \hat{x} = \arg \max_x p(x|y) \\ \hat{\theta} = \arg \max_{\theta} p(\theta|y, \hat{x}) \end{cases} \quad \text{avec } p(x|y) = \int_{\theta} p(x, \theta|y) d\theta \quad (1.35)$$

Mais le calcul théorique des lois conditionnelles $p(\theta|y)$ ou $p(x|y)$ peut être impossible (annexe C). Cette difficulté peut être contournée par l'emploi d'un algorithme itératif calculant successivement \hat{x} et $\hat{\theta}$ (JMAP hiérarchique) :

$$\begin{cases} \hat{\theta} = \arg \max_{\theta} p(\theta|y, \hat{x}) \\ \hat{x} = \arg \max_x p(x|y, \hat{\theta}) \end{cases} \quad (1.36)$$

qui peut ne pas converger vers l'optimum global $(\widehat{x}, \widehat{\theta})$ au sens du MAP (annexe F). L'emploi d'une méthode de *Monte-Carlo*, utilisant les mêmes lois a posteriori $p(\theta|y, \widehat{x})$ et $p(x|y, \widehat{\theta})$ (annexe D), permet alors de trouver un estimateur valable, au sens de la MMSE.

1.6 Influence du bruit

Les données tomographiques sont sujettes à de nombreuses sources de bruit. Les faisceaux de photons sont polychromatiques, d'une certaine largeur. Les photons peuvent être absorbés par la matière, mais aussi dispersés, pour être reçus par d'autres détecteurs. Ces détecteurs, comme la source de rayons X, ne sont pas ponctuels. Les données sont quantifiées. Le modèle :

$$y = f_{\theta}(x) + \epsilon(\theta) \quad (1.37)$$

est imparfait. Dans le domaine médical, la quantité de photons X qui irradie le patient doit être limitée. Les mouvements du patient lors de l'analyse tomographique, les dépassements de la zone d'étude (patient obèse) introduisent des fluctuations dans les mesures. Enfin, le nombre de projections est réduit (en médical, en *contrôle non destructif* industriel) pour des raisons économiques et pratiques.

La modélisation de l'ensemble de ces imprécisions est ardue. Les méthodes probabilistes permettent de contourner la difficulté, en donnant au bruit une loi, plus ou moins réaliste, le caractérisant. Ici, c'est un *bruit additif blanc gaussien* qui a été choisi, de *rapport signal/bruit* (rapport des variances du sinogramme et du bruit ajouté) de l'ordre de $20dB$. Ce bruit est donc supposé indépendant des mesures. Une modélisation plus réaliste du bruit (les mesures sur les détecteurs ayant reçu peu de photons – ie pour des rayons ayant traversé une importante quantité de matière – devraient être plus sensibles au bruit) complexifierait les algorithmes d'estimation (par l'écriture des lois *a posteriori*, lors de la maximisation ou des tirages).

L'ajout d'une erreur *non systématique* indépendante gaussienne (comme somme de lois indépendantes, d'après le théorème de la limite centrale) paraît donc raisonnable.

La modélisation des imprécisions ϵ restera ici l'ajout d'un bruit blanc gaussien

La reconstruction devra donc être réalisée à partir de *mesures incomplètes* (§1.2.4) et *bruitées*. Une *bonne* méthode d'inversion devra en particulier être *robuste au bruit*.

Chapitre 2

Fusion homogène en reconstruction tomographique

2.1 Reconstruction à partir d'un faible nombre de projections

Une reconstruction exacte à partir des données tomographiques seules est impossible (§1.2.4), mais peut-être approchée pour un grand volume de données.

Lorsque celui-ci est plus limité, le caractère *mal posé* du problème d'inversion s'affirme, pour des résultats de reconstruction catastrophiques (FIG 2.1 (2)). L'introduction d'un *a priori* markovien traduisant l'homogénéité du champ à reconstruire (et qui se traduit par l'ajout d'un terme de *régularisation* à sa logvraisemblance [5]) améliore sensiblement les résultats (FIG 2.1 (3)). Le problème reste néanmoins mal posé, et l'introduction de données supplémentaires (z, s) est nécessaire pour aider à la décision.

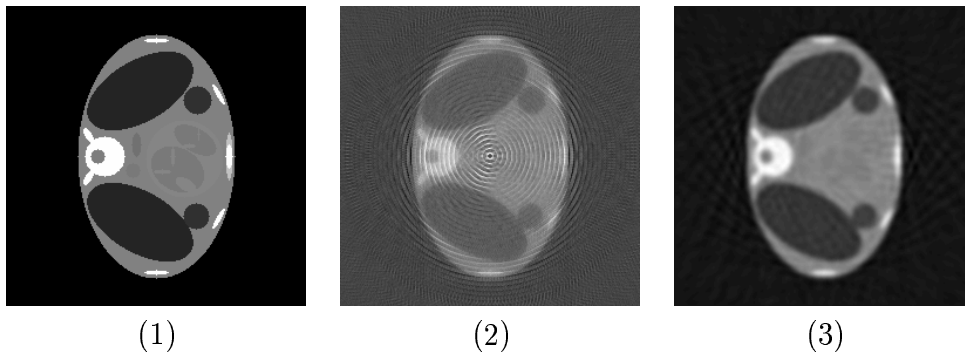


FIG. 2.1 – Fantôme (1), reconstruction à partir des seules données tomographiques (2), avec ajout d'un *a priori* markovien (3)

2.2 Modèle et méthodes d'estimation

Vont être fusionnées des données *radiographiques* (y), *échographiques* (z) et *anatomiques* (s) (FIG.2.2).

Soit : x la fonction d'atténuation de la matière étudiée, à *reconstruire*,
 y les projections tomographiques (mesures de x),
 z les données échographiques (mesures de la réflexivité acoustique),
 s un *a priori* des régions connues de x , qui peut se réduire à une simple silhouette.

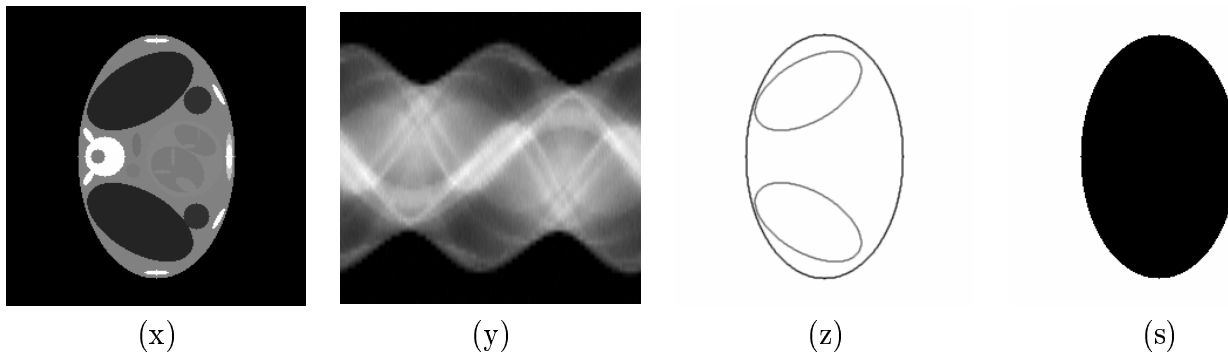


FIG. 2.2 – Fantome (x), sinogramme (y), contours connus (z), régions connues (s)

L'objectif de la méthode présentée consiste à effectuer une reconstruction tomographique, soit estimer x , quitte à estimer d'autres champs pour y aider (variances, contours, ...), à partir des données radiographiques y , acoustiques z et anatomiques s .

Les données radiographiques y sont directement liées à l'image à reconstruire x par le modèle $y = H_1x + \epsilon_1$ ((1.20) et (1.26)).

x exprime la fonction d'atténuation de la matière étudiée, *mais non exactement* sa densité (§1.1.1). De même, les données échographiques z mesurent la réflexivité acoustique de la matière, qui dépend fortement mais *en partie* des variations de la densité de celle-ci. La relation entre mesures échographiques z et fonction d'atténuation x , via la densité, n'est donc pas évidente. Néanmoins, on suppose possible dans ce modèle la mesure directe de z^1 à partir de x^2 , soit $z = H_2x + \epsilon_2$.

L'estimation sera bayésienne, par minimisation d'un critère (1.29) ou à l'aide d'une méthode de Monte-Carlo (annexe D et (1.30)).

Tous les champs (images x, z, s et le sinogramme y) seront représentés par la suite sous forme vectorielle (§1.3.2).

2.2.1 Modèle

Les données (y, z, s) sont toutes liées à l'image à reconstruire x par le modèle :

$$\begin{cases} y = H_1x + \epsilon_1 \\ z = H_2x + \epsilon_2 \\ s = H_3x + \epsilon_3 \end{cases} \quad \epsilon_i \sim \mathcal{N}(0, \sigma_i^2 I_{N_i}) \quad (2.1)$$

¹si l'on suppose possible l'écriture d'une relation bijective entre la fonction d'atténuation x et la densité de la matière étudiée, et que l'extraction de données échographiques z peut être modélisée par un filtrage (opération linéaire) sur cette même densité, on admet la possibilité d'une mesure directe de z à partir de x .

²si les contours (binaires) de la carte de densité s'identifient à ceux de la fonction d'atténuation (x), un traitement des données échographiques devrait pouvoir apporter des résultats proches de ceux utilisés en simulation (et plus ou moins complets).

où : H_1 est la matrice des projections (§1.3.2) qui simule le fonctionnement d'un tomographe,
 H_2 est un opérateur qui simule le fonctionnement d'un échographe,
 H_3 est un masque sur le champ x .

Une justification sur le caractère gaussien indépendant des bruits ajoutés aux données tomographiques a été donnée plus haut (§1.6), et on suppose qu'il peut s'étendre aux mesures échographiques et anatomiques.

Pour faciliter à la décision, un *a priori* d'homogénéité sur x est ajouté, modélisé (annexe A) par une loi markovienne d'énergie $(\alpha/2)U(x)$.

Le coefficient α sera à estimer, tout comme les variances σ_i^2 des lois du bruit. Soit $\theta = (\theta_1, \theta_2, \theta_3, \theta_4) = (1/\sigma_1^2, 1/\sigma_2^2, 1/\sigma_3^2, \alpha)$ l'ensemble des hyperparamètres.

Seront estimés dans cette méthode de fusion homogène
l'image à reconstruire \hat{x} et les hyperparamètres $\hat{\theta}_i$

D'un peu plus près...

H_1 est la matrice de projection définie au §1.3.2. La multiplication matricielle H_1x est effectuée directement en pratique.

H_3 a été défini comme un masque de confiance sur le champ x , soit :

$$H_3 = \text{diag}(\mu_j) \quad , \quad \mu_j \in \{0, 1\} \quad (2.2)$$

avec $\mu_j = 1$ pour une pleine confiance en la valeur s_j pour x_j , et $\mu_j = 0$ pour aucun *a priori* sur la valeur de x_j .

Dans le cas où l'*a priori* sur x est introduit avec une incertitude ($\mu_j \in [0, 1]$, les valeurs intermédiaires pour une connaissance incertaine), on peut écrire :

$$s = x + \epsilon_3 \quad \text{avec} \quad \epsilon_3 \sim \mathcal{N}(0, \sigma_3^2 H_3^{-1}) \quad \text{et} \quad H_3 = \text{diag}(\mu_j) \quad (2.3)$$

Le premier modèle sera utilisé par la suite.

Les informations échographiques portent *plus ou moins* sur les contours directionnels. On suppose ici que :

De simples extractions de contours apportent une
information *équivalente* à une étude échographique

Ce choix peu réaliste de H_2 n'a alors pas d'importance lors des simulations de fusion sur des fantomes. Mais une modélisation plus avancée serait intéressante pour tester la méthode sur des données réelles.

Afin de s'approcher au maximum de la réalité échographique, nous décomposerons H_2 en quatre opérateurs directionnels de dérivation discrète, pour les directions horizontales (H_{21}), verticales (H_{22}) et diagonales (H_{23} et H_{24}). En décomposant de même les données échographiques z (z_i) et leurs bruits associés, toujours gaussiens centrés et indépendants (§1.6), ϵ_2 (ϵ_{2i} de variances $\sigma_{2i}^2 = 1/\theta_{2i}$), nous avons ainsi les quatres relations :

$$z_i = H_{2i}x + \epsilon_{2i} \quad , \quad i \in \{1, 2, 3, 4\} \quad (2.4)$$

Pour un choix des paramètres θ_{2i} égaux entre eux (bruits ϵ_{2i} de même variances), ces quatres équations se résument à :

$$z = \left(\sum_i H_{2i} \right) x + \epsilon_2 \quad (2.5)$$

Les choix des opérateurs de dérivation H_{2i} (E.2) font de cette opération une extraction omnidirectionnelle de contours.

Enfin, la loi markovienne du champ x est un *a priori* d'homogénéité, et son énergie se met par exemple sous la forme (A.10), soit :

$$U(x) = x^\top M x \quad (2.6)$$

2.2.2 Ecriture des lois *a posteriori*

L'estimation de (θ, x) par MAP encourage à une *maximisation disjointe* (annexe C et §1.5.4), soit remplacer l'étude de la loi de $(x, \theta|y, z, s)$ par celles, *non équivalentes* (annexe F), de $(x|y, z, s, \theta)$ et $(\theta|y, z, s, x)$. Les méthodes de Monte-Carlo font intervenir ces mêmes lois lors des tirages. Il faut donc chercher à écrire, dans les deux cas, $p(x|y, z, s, \theta)$ et $p(\theta|y, z, s, x)$.

Calcul de la loi *a posteriori* de l'image à reconstruire

Elle s'écrit (annexe C et §1.5.3) :

$$\begin{aligned} p(x|y, z, s, \theta) &\propto p(y|x, z, s, \theta).p(x|z, s, \theta) \\ &\propto p(y|x, \theta).p(x|z, s, \theta) \\ &\propto p(y|x, \theta).p(z|x, \theta).p(s|x, \theta).p(x|\theta) \\ &\propto p(y|x, \theta).p(z|x, \theta).p(s|x, \theta).p(x) \\ &\propto p(y|x, \theta). \left(\prod_{i=1}^4 p(z_i|x, \theta) \right) .p(s|x, \theta).p(x) \end{aligned} \quad (2.7)$$

Ceci, avec la loi *a priori* markovienne pour x et les lois du bruit de (2.1) :

$$\begin{cases} y|x, \theta \sim \mathcal{N}(H_1 x, \sigma_1^2 I_{N_y}) \\ z_i|x, \theta \sim \mathcal{N}(H_{2i} x, \sigma_{2i}^2 I_{N_x}) \quad , i \in \{1, 2, 3, 4\} \\ s|x, \theta \sim \mathcal{N}(H_3 x, \sigma_3^2 I_{N_x}) \\ p(x) \propto \exp \left(-\frac{\alpha}{2} x^\top M x \right) \end{cases} \quad (2.8)$$

Soit la probabilité de x *a posteriori* :

$$p(x|y, z, s, \theta) \propto \exp \left[-\frac{1}{2} \left(\theta_1 \|y - H_1 x\|^2 + \sum_{i=1}^4 \theta_{2i} \|z_i - H_{2i} x\|^2 + \theta_3 \|s - H_3 x\|^2 + \theta_4 x^\top M x \right) \right] \quad (2.9)$$

Calcul de la loi *a posteriori* des hyperparamètres

Elle s'écrit :

$$\begin{aligned} p(\theta|y, z, s, x) &\propto p(y|\theta, z, s, x).p(\theta|z, s, x) \\ &\propto p(y|x, \theta).p(\theta|z, s, x) \\ &\propto p(y|x, \theta).p(z|x, \theta).p(s|x, \theta).p(\theta|x) \\ &\propto p(y|x, \theta). \left(\prod_{i=1}^4 p(z_i|x, \theta) \right) .p(s|x, \theta).p(x|\theta).p(\theta) \end{aligned} \quad (2.10)$$

Puis, en redécomposant $\theta = (\theta_1, \theta_{21}, \theta_{22}, \theta_{23}, \theta_{24}, \theta_3, \theta_4)$, les θ_i étant indépendants entre eux, on trouve :

$$\begin{cases} p(\theta_1|y, z, s, x) \propto p(y|x, \theta_1).p(\theta_1) \\ p(\theta_{2i}|y, z, s, x) \propto p(z_i|x, \theta_{2i}).p(\theta_{2i}) \quad , i \in \{1, 2, 3, 4\} \\ p(\theta_3|y, z, s, x) \propto p(s|x, \theta_3).p(\theta_3) \\ p(\theta_4|y, z, s, x) \propto p(x|\theta_4).p(\theta_4) \end{cases} \quad (2.11)$$

En supposant les lois *a priori* des variables aléatoires θ_i gamma (annexe C), de paramètres (a_i, b_i) fixés :

$$p(\theta_i) \propto \theta_i^{a_i-1} \exp(-b_i\theta_i) \quad (2.12)$$

les probabilités de ces mêmes variables, mais *a posteriori*, sont aussi des lois gamma, de paramètres s'exprimant en fonction des erreurs d'estimation et de l'énergie du champ x :

$$\begin{cases} p(\theta_1|y, z, s, x) \propto \theta_1^{a_1+N_y/2-1} \exp[-(b_1 + \|y - H_1x\|^2/2)\theta_1] \\ p(\theta_{2i}|y, z, s, x) \propto \theta_{2i}^{a_{2i}+N_x/2-1} \exp[-(b_{2i} + \|z_i - H_{2i}x\|^2/2)\theta_{2i}] \quad , i \in \{1, 2, 3, 4\} \\ p(\theta_3|y, z, s, x) \propto \theta_3^{a_3+N_x/2-1} \exp[-(b_3 + \|s - H_3x\|^2/2)\theta_3] \\ p(\theta_4|y, z, s, x) \propto \theta_4^{a_4+N_x/2-1} \exp[-(b_4 + x^T Mx/2)\theta_4] \end{cases} \quad (2.13)$$

2.2.3 Estimation par MAP

Paramètres θ et image à reconstruire x vont tout deux être estimés par MAP. L'estimation de θ se fera analytiquement, celle de x en minimisant un critère par un algorithme de descente [5]. L'algorithme utilisé sera celui de (1.36), par estimation successive de \hat{x} et $\hat{\theta}$ (§1.5.4).

Optimum des hyperparamètres

Le MAP des lois des θ_i permet de trouver leur optimum (annexe C) :

$$\begin{cases} \hat{\theta}_1 = (a_1 + N_y/2 - 1)/(b_1 + \|y - H_1x\|^2/2) \\ \hat{\theta}_{2i} = (a_{2i} + N_x/2 - 1)/(b_{2i} + \|z_i - H_{2i}x\|^2/2) \quad , i \in \{1, 2, 3, 4\} \\ \hat{\theta}_3 = (a_3 + N_x/2 - 1)/(b_3 + \|s - H_3x\|^2/2) \\ \hat{\theta}_4 = (a_4 + N_x/2 - 1)/(b_4 + x^T Mx/2) \end{cases} \quad (2.14)$$

Reste à estimer l'optimum de l'image à reconstruire, \hat{x} , connaissant ces valeurs optimales $\hat{\theta}_i$.

Critère à minimiser

La nature des lois (gaussiennes et markoviennes (2.8) permet d'écrire simplement le logarithme de la probabilité *a posteriori* (2.7) de l'image à reconstruire. A une constante additive près :

$$\begin{aligned} J(x) &\equiv -2 \log \left(p(x|y, z, s, \theta) \right) \\ &= \theta_1 \|y - H_1x\|^2 + \sum_{i=1}^4 \theta_{2i} \|z_i - H_{2i}x\|^2 + \theta_3 \|s - H_3x\|^2 + \theta_4 x^T Mx \\ &= \sum_i \theta_i J_i(x) \text{ en écrivant } \theta_2 \|z - H_2x\|^2 = \sum_{i=1}^4 \theta_{2i} \|z_i - H_{2i}x\|^2 \end{aligned} \quad (2.15)$$

Maximiser la loi *a posteriori* $p(x|y, z, s, \theta)$, c'est minimiser le critère $J(x)$, qui s'écrit comme somme de 4 sous-critères $J_i(x)$ (mesures tomographiques, mesures échographiques, *a priori* des régions connues et homogénéité de l'image à reconstruire) pondérés par les paramètres θ_i

Un bon choix des variances σ_i et du coefficient α paraît donc déterminant dans la procédure de reconstruction.

Le critère étudié est quadratique (2.15), et son minimum peut-être calculé analytiquement :

$$J(x) = x^\top Ax + Bx + C \quad \text{d'optimum} \quad \hat{x} = -\frac{1}{2}A^{-1}B \quad (2.16)$$

Mais la matrice A est ici *mal conditionnée*, soit d'inverse instable et inutilisable en pratique (et, même si l'inversion directe était possible, son coût serait trop grand) [5].

L'optimum sera alors calculé différemment, ici itérativement, par un algorithme du gradient.

Cette méthode a l'avantage de pouvoir être aussi applicable dans le cas de lois du bruit plus générales (gaussiennes généralisées), d'*a priori* markoviens non quadratiques, d'opérateurs de dérivation H_i non linéaires, ...

Gradient du critère

Afin d'utiliser les outils de minimisation, le gradient du critère $J(x)$ va devoir être calculé analytiquement. En utilisant les gradients matriciels (annexe B), on peut écrire :

$$\nabla J(x) = -2\theta_1 H_1^\top (y - H_1 x) - 2 \sum_{i=1}^4 \theta_{2i} H_{2i}^\top (z_i - H_{2i} x) + 2\theta_3 \sum_j \mu_j (x_j - s_j) + \theta_4 (Mx + M^\top x) \quad (2.17)$$

et sachant les H_{2i} et M symétriques (annexe E), on trouve :

$$\nabla J(x) = 2 \left[\theta_1 H_1^\top (H_1 x - y) + \sum_{i=1}^4 \theta_{2i} H_{2i} (H_{2i} x - z_i) + \theta_3 H_3 (x - s) + \theta_4 Mx \right] \quad (2.18)$$

où : $H_1^\top (H_1 x - y)$ est la rétroprojection de la différence des données tomographiques avec la projection de l'estimée,

$H_{2i} (H_{2i} x - z_i)$ est l'extraction des contours de la différence des données échographiques avec ceux de l'estimée,

$H_3 (x - s)$ est la différence de l'*a priori* avec l'estimée, sur les régions connues,

Mx est la dérivée discrète de l'estimée.

Algorithme

C'est l'application directe de (1.36) :

1. initialisation $x^{(0)}$ de l'image à reconstruire comme rétroprojection des données tomographiques (§1.4.3),
2. jusqu'à convergence, estimer :
 - les $\hat{\theta}_i$ simplement calculés par (2.14),
 - \hat{x} en appliquant un algorithme du gradient.

2.2.4 Estimation par méthode de Monte-Carlo

Image à reconstruire et paramètres (x, θ) vont être estimés par MMSE. Des échantillons $(x^{(t)}, \theta^{(t)})$ vont être générés selon une méthode de Monte-Carlo, et suivront asymptotiquement la loi *a posteriori* $p(x, \theta|y, z, s)$. La moyenne arithmétique des derniers :

$$\widehat{(x, \theta)} = \frac{1}{\Delta T} \sum_{t=T_0}^{T_0+\Delta T} (x^{(t)}, \theta^{(t)}) \quad (2.19)$$

sera alors une estimation de l'espérance du processus $(x^{(t)}, \theta^{(t)})$, soit de la moyenne de sa loi $p(x, \theta|y, z, s)$, et donc une estimation de (x, θ) au sens de la MMSE.

Le choix de la méthode (Gibbs ou hybride, annexe D) dépendra de la forme du bruit ajouté aux mesures.

Méthode choisie

Soit $x = (x_k)$ l'image à reconstruire, $\theta = (\theta_i)$ les hyperparamètres à estimer. Il faut chercher à identifier (algorithme de Gibbs) ou écrire (hybride) les lois :

$$\begin{cases} p(\theta_i|x, \theta_{j \neq i}, y, z, s) \\ p(x_k|x_{l \neq k}, \theta, y, z, s) \end{cases} \quad (2.20)$$

L'échantillonneur de Gibbs est plus efficace que sa version hybride (le taux d'acceptation est égal à 1 lorsque la loi instrumentale s'identifie à la loi étudiée). Ce n'est donc pas le simple calcul (à une constante multiplicative près) des probabilités (2.20) qui est recherché, mais leur identification (nature et paramètres), afin d'effectuer directement des tirages à partir de celles-ci. Cette identification est possible ici.

Tirage des hyperparamètres

L'hypothèse d'indépendance des hyperparamètres entre eux permet d'écrire $p(\theta_i|x, \theta_{j \neq i}, y, z, s) = p(\theta_i|y, z, s, x)$, qui est, d'après (2.13), une loi gamma. Ainsi :

Les tirages des hyperparamètres θ_i se feront directement suivant des lois gamma

$$\begin{cases} \theta_1|x, \theta_{j \neq 1}, y, z, s \sim \gamma(a_1 + N_y/2, b_1 + \|y - H_1x\|^2/2) \\ \theta_{2i}|x, \theta_{j \neq 2i}, y, z, s \sim \gamma(a_{2i} + N_x/2, b_{2i} + \|z_i - H_{2i}x\|^2/2) \\ \theta_3|x, \theta_{j \neq 3}, y, z, s \sim \gamma(a_3 + N_x/2, b_3 + \|s - H_3x\|^2/2) \\ \theta_4|x, \theta_{j \neq 4}, y, z, s \sim \gamma(a_4 + N_x/2, b_4 + x^T M x/2) \end{cases} \quad , i \in \{1, 2, 3, 4\} \quad (2.21)$$

Tirage des pixels de l'image

La probabilité *a posteriori* d'un pixel x_k , conditionnellement aux autres pixels de l'image, suit une loi gaussienne, de moyenne et variance connues. Ainsi :

Les tirages des pixels x_k de l'image à reconstruire se feront directement suivant des lois gaussiennes

Ceci, dans le cas de bruits gaussiens, et d'un *a priori* markovien quadratique (2.6). Dans le cas contraire, on cherchera seulement à exprimer les probabilités conditionnelles des pixels (à une constante multiplicative près). Et on pourra prendre comme loi instrumentale celle calculée ci-dessous dans le cas gaussien (avec quelques précautions³), et utiliser l'algorithme hybride. Ici, dans le cas gaussien, on peut écrire à partir de (2.9), sachant $M = N.N$ et $N = N^T$:

$$\begin{aligned} J(x) &\equiv -\log(p(x|y, z, s, \theta)) \\ &= \frac{1}{2} [\theta_1 \|y - H_1 x\|^2 + \sum_{i=1}^4 \theta_{2i} \|z_i - H_{2i} x\|^2 + \theta_3 \|s - H_3 x\|^2 + \theta_4 \|Nx\|^2] \\ &= \frac{1}{2} \|Y - Hx\|^2 \end{aligned} \quad (2.22)$$

avec :

$$Y = \begin{pmatrix} \sqrt{\theta_1} \cdot y \\ \sqrt{\theta_2} \cdot z \\ \sqrt{\theta_3} \cdot s \\ \sqrt{\theta_4} \cdot 0 \end{pmatrix} \quad H = \begin{pmatrix} \sqrt{\theta_1} \cdot H_1 \\ \sqrt{\theta_2} \cdot H_2 \\ \sqrt{\theta_3} \cdot H_3 \\ \sqrt{\theta_4} \cdot N \end{pmatrix} \quad (2.23)$$

$J(x)$ est quadratique en x , donc en x_k , $\forall k$. Et comme :

$$\begin{aligned} p(x_k | x_{l \neq k}, \theta, y, z, s) &\propto p(x_k, x_{l \neq k}, \theta, y, z, s) / p(x_{l \neq k}, \theta, y, z, s) \\ &\propto p(x | y, z, s, \theta) \end{aligned} \quad (2.24)$$

la loi de $x_k | x_{l \neq k}, \theta, y, z, s$ sera gaussienne, de variance et moyenne :

$$\begin{cases} \sigma_k^2 = 1 / \left(\frac{\partial^2}{\partial x_k^2} J(x) \right) \\ m_k = x_k \text{ tel que } \left. \frac{\partial}{\partial x_k} J(x) \right|_{m_k} = 0 \end{cases} \quad (2.25)$$

à la seule condition que :

$$\frac{\partial^2}{\partial x_k^2} J(x) \neq 0 \quad (2.26)$$

On remarque que le vecteur $x | y, z, s, \theta$ n'est lui-même pas nécessairement gaussien, car

$$\frac{\partial^2}{\partial x^2} J(x) = H^T H \quad (2.27)$$

peut ne pas être inversible (mal conditionnée ou singulière)[5]. La matrice de covariance de la loi gaussienne ne peut alors être définie.

Calcul de la variance

On notera $C_k(P)$ la k^{eme} colonne de la matrice P et $L_k(P)$ sa k^{eme} ligne. Le gradient de $J(x)$ est le vecteur :

$$\frac{\partial}{\partial x} J(x) = -H^T (Y - Hx) \quad (2.28)$$

³le maximum du produit de deux gaussiennes ne s'identifie pas à celui du produit de deux exponentielles généralisées, à variances et moyennes inchangées.

de k^{eme} composante :

$$\begin{aligned}\frac{\partial}{\partial x_k} J(x) &= -L_k(H^\top)(Y - Hx) \\ &= -C_k(H)^\top(Y - Hx)\end{aligned}\quad (2.29)$$

Le hessien de $J(x)$ est la matrice :

$$\frac{\partial^2}{\partial x^2} J(x) = H^\top H \quad (2.30)$$

d'élément (k,k) :

$$\begin{aligned}\frac{\partial^2}{\partial x_k^2} J(x) &= L_k(H^\top)C_k(H) \\ &= \|C_k(H)\|^2 \\ &= \sum_i \theta_i \|C_k(H_i)\|^2 \\ &\equiv 1/\sigma_k^2\end{aligned}\quad (2.31)$$

l'inverse de la variance de la loi $x_k|x_{l \neq k}, \theta, y, z, s$, indépendante de x , et qui peut donc être précalculée.

Calcul de la moyenne

(2.25) et (2.29) permettent d'écrire, en notant $x = x_k \cdot e_k + \bar{x}$:

$$\begin{aligned}\frac{\partial}{\partial x_k} J(x) &= -C_k(H)^\top(Y - H(x_k \cdot e_k + \bar{x})) \\ &= -C_k(H)^\top(Y - H\bar{x}) + x_k \|C_k(H)\|^2\end{aligned}\quad (2.32)$$

qui s'annule en :

$$\begin{aligned}m_k &= \sigma_k^2 C_k(H)^\top(Y - H\bar{x}) \\ &= \sigma_k^2 \sum_i \theta_i C_k(H_i)^\top(y_i - H_i \bar{x})\end{aligned}\quad (2.33)$$

la moyenne de la loi $x_k|x_{l \neq k}, \theta, y, z, s$. *Seul* le calcul $C_k(H)^\top(Y - H\bar{x})$ sera à effectuer pour chaque tirage.

Algorithme

Chaque échantillon $(x, \theta)^{(t)}$ est généré par l'algorithme de Gibbs (D.3). Une moyenne arithmétique des derniers (2.19) nous donne alors une estimation des hyperparamètres et de l'image à reconstruire :

1. initialisation $x^{(0)}$ de l'image à reconstruire comme rétroprojection des données tomographiques (§1.4.3),
2. précalculs : écriture des matrices (creuses) H_j sous forme compactes (colonne par colonne), calculs des variances σ_k^2 (2.31),
3. jusqu'à $T_0 + \Delta T$ itérations, générer :
 - chacune des variances $1/\theta_i$ à partir des lois gamma (2.21),
 - chaque pixel x_k selon la loi conditionnelle $\mathcal{N}(m_k, \sigma_k^2)$, pour obtenir un échantillon $x^{(t)}$.
4. calculer \hat{x} et les $\hat{\theta}_i$, moyennes des ΔT derniers échantillons correspondants.

2.3 Simulations

Elles ont toutes été effectuées à partir des contours et régions connues de la FIG.2.2, sans utiliser l'*a priori* markovien⁴. Les données échographiques ont été introduites via les contours omnidirectionnels seuls.

La qualité de la reconstruction est ici évaluée par le calcul des distances des images reconstruites à l'original :

$$d_1(x_{estimate}, x_0) = \frac{\sum_k |x_{estimate}(k) - x_0(k)|}{\sum_k |x_0(k)|} \quad d_2(x_{estimate}, x_0) = \frac{\sum_k (x_{estimate}(k) - x_0(k))^2}{\sum_k x_0(k)^2} \quad (2.34)$$

2.3.1 MAP VS MCMC

Coût de calculs

La génération d'un échantillon (MCMC) est équivalente en coût à une itération dans l'algorithme du gradient (MAP)

En effet, l'échantillonneur de Gibbs requiert, à chaque tirage, le calcul des moyennes m_k (2.33) pour chacun des pixels k , soit, pour toute l'image, l'évaluation matricielle $H^T(Y - Hx)$. Et c'est ce même calcul qui intervient dans l'évaluation du gradient (2.18) à chaque itération.

Cette équivalence en nombre d'opérations n'implique pas celle en temps de calcul, sous Matlab (où "les boucles *for* sont à éviter comme la peste"⁵).

Paramètres dans les simulations

Les simulations présentées ont ainsi été réalisées à partir d'images de basse résolution (64×64). Les calculs MCMC de plusieurs heures devraient, dans un langage moins évolué que Matlab, se réduire aux quelques secondes de ceux obtenus par MAP.

Chaque simulation a nécessité un nombre d'itérations/échantillons comparable (500), et ont été effectuées à partir des données de la FIG.2.2, à résolution inférieure.

Les premiers résultats ont été effectués à partir de données (FIG.2.3) non bruitées (FIG.2.4), puis en ajoutant sur les données tomographiques un bruit de 20dB (FIG.2.5) et 10dB (FIG.2.6).

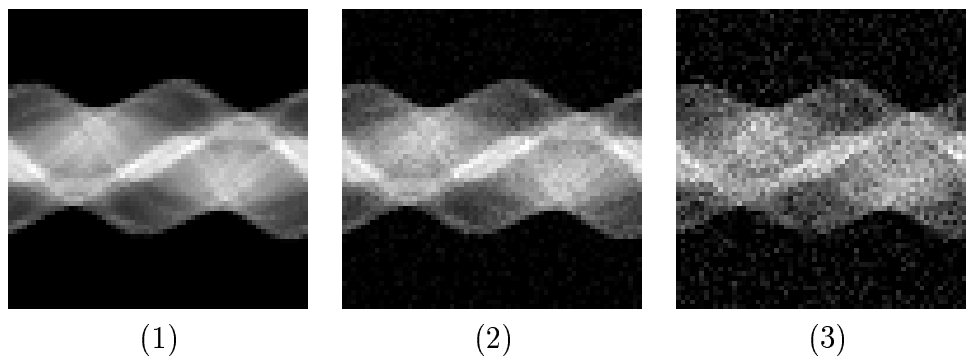


FIG. 2.3 – Données tomographiques non bruitées (1), avec un bruit additif gaussien de rapport S/N de 20dB (2), 10dB (3)

⁴l'introduction des contours et régions connues régularise déjà l'inversion de $H^T H$. Le MAP est alors MV.

⁵Mahieddine Mehdi Ichir

Estimation des hyperparamètres

Les méthodes de Monte-Carlo ont l'avantage sur le MAP de permettre, avec la reconstruction tomographique, une estimation fiable des hyperparamètres θ_i . Ils s'écrivent, à l'optimum :

$$\hat{\theta} = \frac{a_0 + N/2 - 1}{b_0 + \|y - Hx\|^2/2} \quad \text{où } N \text{ est la dimension de } y \quad (2.35)$$

soit :

$$\begin{aligned} \sigma^2 &\simeq \frac{1}{N} \|y - Hx\|^2 \\ &\simeq \frac{1}{N} \sum_{k=1}^N (y_k - (Hx)_k)^2 \end{aligned} \quad (2.36)$$

qui est bien une estimation (à partir de N échantillons) de la variance du bruit blanc ajouté sur les mesures.

Un choix incorrect des variances ne permet pas la reconstruction [1]. Leur estimation permet ainsi d'éviter, pour les reconstructions futures, de leur difficiles choix empiriques.

Estimation du rapport (S/N)

Une évaluation du rapport (S/N) des données tomographiques peut être réalisée, ceci à partir de l'estimation de la variance σ_1 . Le rapport signal sur bruit devrait être calculé connaissant la variance du signal (des mesures tomographiques) non bruité :

$$(\widehat{S/N})_0 = 10 \log \left(\frac{\sigma_{\text{sinogramme non bruité}}}{\sigma_1} \right) \quad (2.37)$$

et est ici aussi calculé à partir du signal reçu :

$$(\widehat{S/N})_1 = 10 \log \left(\frac{\sigma_{\text{sinogramme bruité}}}{\sigma_1} \right) \quad (2.38)$$

où les variances $\sigma_{\text{sinogramme non bruité}}$ et $\sigma_{\text{sinogramme bruité}}$ sont estimées comme écarts quadratiques moyens des valeurs des sinogrammes respectifs.

Les résultats montrent une estimation proche du rapport dans les deux cas (TAB. 2.1 et 2.2), qui est donc possible en pratique (via le calcul de $(\widehat{S/N})_1$).

| (S/N) (dB) | d_1 | d_2 | $\hat{\theta}_1$ (10^{-3}) | $\hat{\sigma}_1^2$ (10^3) | $(\widehat{S/N})_0$ (dB) | $(\widehat{S/N})_1$ (dB) |
|--------------|-------|-------|--------------------------------|-------------------------------|--------------------------|--------------------------|
| ∞ | 0.11 | 0.036 | 1.10 | 0.9 | 26.7 | 26.7 |
| 20 | 0.12 | 0.042 | 0.42 | 2.3 | 22.0 | 21.8 |
| 10 | 0.75 | 0.242 | 0.04 | 20.7 | 11.1 | 10.8 |

TAB. 2.1 – Résumé des résultats obtenus par MAP hiérarchique

Les reconstructions tomographiques sont imparfaites, ce qui explique le caractère borné des estimations du rapport (S/N).

Dans le cas de signaux bruités, les estimées sont légèrement biaisées par valeur supérieure, par introduction des données échographiques (z) et anatomiques (s) qui améliorent le rapport signal sur bruit.

| (S/N) (dB) | d_1 | d_2 | $\hat{\theta}_1$ (10^{-3}) | $\hat{\sigma}_1^2$ (10^3) | $(\widehat{S/N})_0$ (dB) | $(\widehat{S/N})_1$ (dB) |
|--------------|-------|-------|--------------------------------|-------------------------------|--------------------------|--------------------------|
| ∞ | 0.10 | 0.028 | 1.29 | 0.8 | 27.5 | 27.5 |
| 20 | 0.13 | 0.038 | 0.45 | 2.2 | 22.2 | 22.0 |
| 10 | 0.16 | 0.055 | 0.07 | 14.6 | 12.8 | 12.5 |

TAB. 2.2 – Résumé des résultats obtenus par méthode de Monte-Carlo (MMSE)

Confiance sur la reconstruction

Les méthodes de Monte-Carlo permettent en outre d'évaluer la confiance à attribuer à l'image reconstruite, *point par point*. En effet, un calcul analogue au (2.19) permet d'estimer la variance des échantillons utilisés pour l'estimation :

$$\widehat{\sigma}_{x_k}^2 = \frac{1}{\Delta T} \sum_{t=T_0}^{T_0+\Delta T} (x_k^{(t)} - \hat{x}_k)^2 \quad \text{pour tout pixel } k \quad (2.39)$$

qui pourrait être utilisée par exemple en imagerie médicale, pour aider au diagnostic. Dans l'algorithme présenté, les pixels sont tirés selon des lois gaussiennes, de variance commune (ne dépendant que des θ_i (2.31)). La carte de variance estimée (2.39) se réduit alors à une simple constante pour toute l'image.

Validité des estimateurs

Les algorithmes du gradient permettent une excellente estimation (au sens du MAP), pour une loi à maximiser *concave*, et n'assurent pas l'atteinte du maximum *global* dans le cas contraire (annexe F). Les méthodes de Monte-Carlo, en revanche, garantissent *asymptotiquement* (et théoriquement) une estimation correcte (au sens de la MMSE ici) pour une loi *quelconque*. Une reconstruction seule, soit sans évaluation des variances du bruit, n'a donc pas d'intérêt par les méthodes de Monte-Carlo pour le modèle quadratique présenté, puisque la loi *a posteriori* $p(x|y, z, s, \theta)$ est concave. En recherchant les variances σ_k^2 , il faut estimer (x, θ) , de loi *a posteriori* $p(x, \theta|y, z, s)$ non nécessairement concave⁶, et les méthodes de Monte-Carlo peuvent devenir intéressantes.

Sensibilité au bruit

Les résultats obtenus (FIG.2.4, 2.5 et 2.6) par MAP, une fois un bruit ajouté aux mesures, oscillent (ces oscillations peuvent s'expliquer par l'existence d'un plateau dans la loi *a posteriori* de x), et ne permettent pas d'estimer les variances. L'estimation de l'image à reconstruire devient elle aussi difficile.

L'estimation par méthodes de Monte-Carlo converge, autant pour les hyperparamètres que les reconstructions, quelle que soit l'amplitude du bruit ajouté. Un point de plus pour Monte-Carlo, car le bruit fait parti de l'étude tomographique (§1.6).

⁶la concavité des lois $p(x|y, z, s, \theta)$ et $p(\theta|y, z, s, x)$ n'assure pas celle de la loi jointe $p(x, \theta|y, z, s)$ (annexe F)

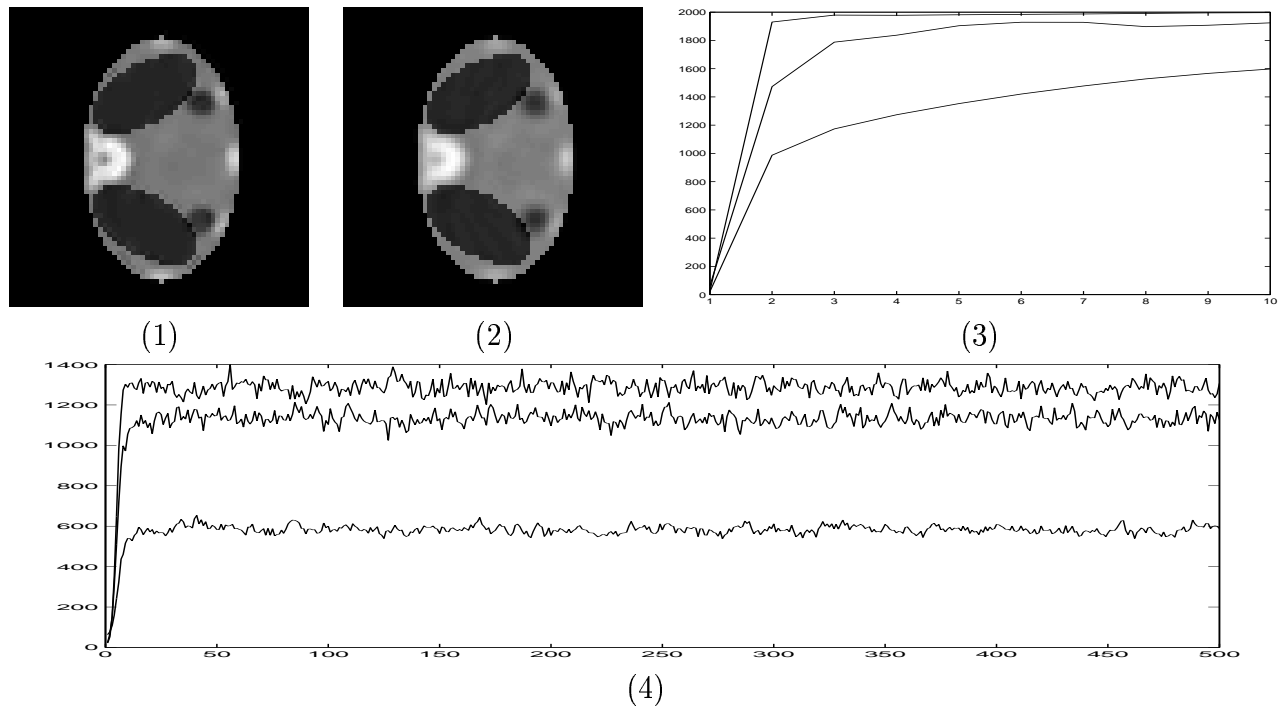


FIG. 2.4 – Image reconstruite par MAP sans bruit ajouté (1), par MCMC (2), hyperparamètres θ_i estimés par MAP (3), par MCMC (4)

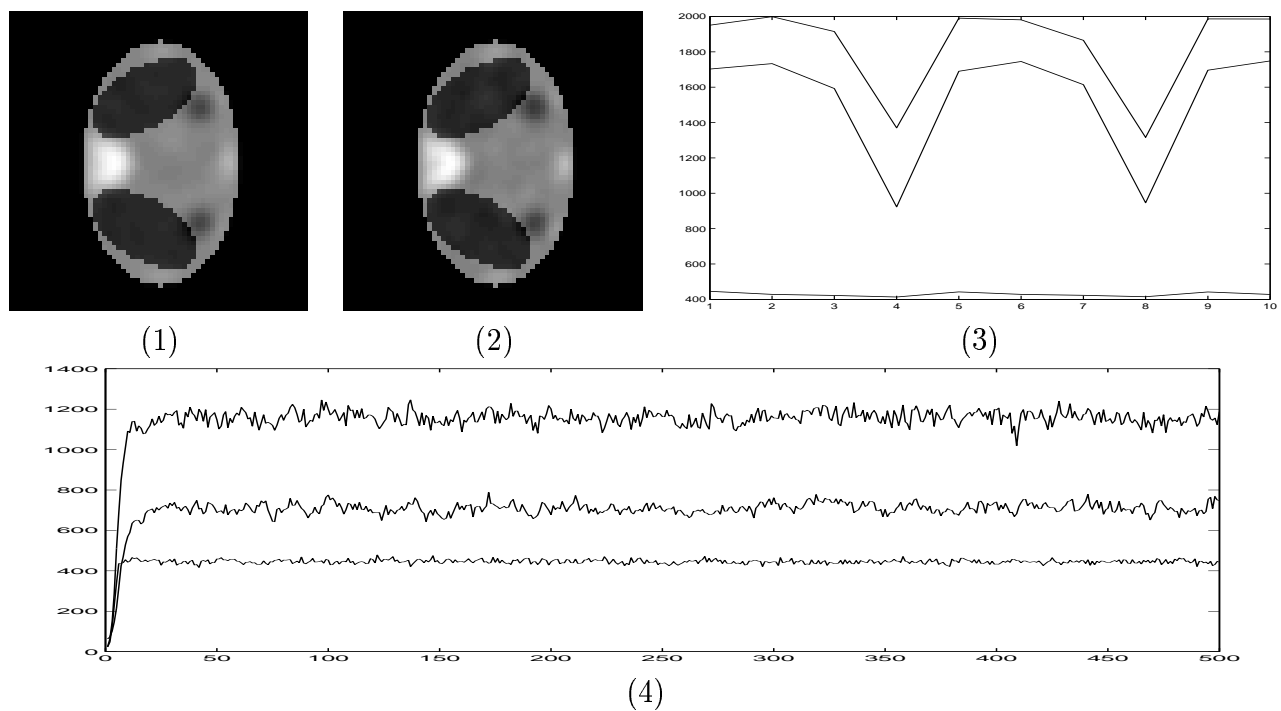


FIG. 2.5 – Image reconstruite par MAP avec un rapport S/N de 20dB sur les données tomographiques (1), par MCMC (2), hyperparamètres θ_i estimés par MAP (3), par MCMC (4)

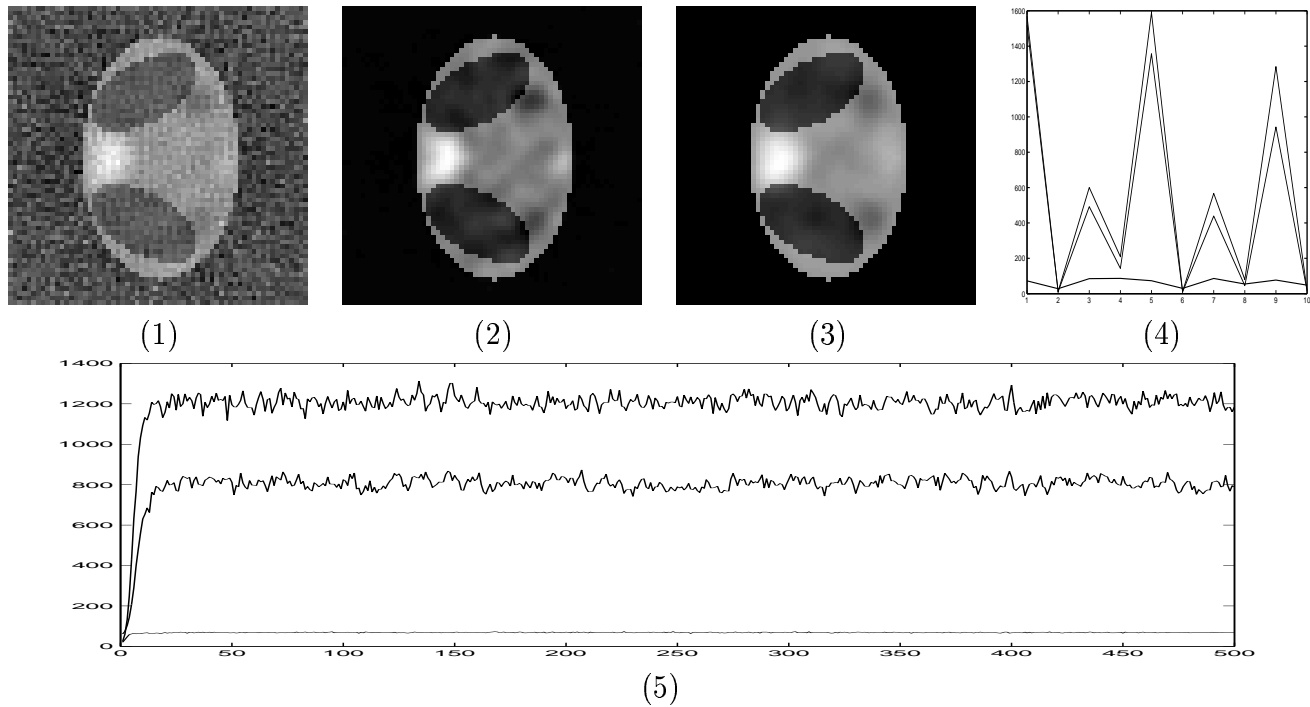


FIG. 2.6 – Images reconstruites par MAP avec un rapport S/N de 10dB sur les données tomographiques (1-2), par MCMC (3), hyperparamètres θ_i estimés par MAP (4), par MCMC (5)

2.3.2 Algorithmes hybrides

ICM

Cette méthode se situe entre le MAP et Monte-Carlo. Elle consiste à appliquer l'échantillonneur (de Gibbs ici), en remplaçant chacun des tirages (selon les lois conditionnelles (2.20) par le maximum des lois correspondantes (m_k (2.33) et $\hat{\theta}_k$ (2.14)). Elle peut être vue comme un MAP hiérarchique, pixel par pixel.

Sa convergence est plus rapide que celle de l'échantillonneur avec tirage, mais sans garantie d'aboutir à une solution valable au sens du MAP (annexe F).

MAP-MCMC

Les hyperparamètres sont générés par échantillonnage, l'image reconstruite par MAP, ou inversement.

La valeur initiale de l'image à reconstruire dans l'échantillonneur MCMC peut-être obtenue par MAP. La convergence, qui est assurée quelle que soit l'initialisation, sera alors plus rapide. Néanmoins, une rétroprojection semble suffire en pratique (convergence après 20 tirages, FIG.2.4, 2.5 et 2.6).

Estimation des données échographiques manquantes

Les mesures échographiques fournissent une partie des contours de l'estimée x , qui aident grandement à la décision (FIG.2.7). Dans le cas du MAP, l'estimation des contours manquants (par *fusion* des contours donnés initialement par les mesures échographiques, et ceux extraits de

l'estimée à une itération donnée) devrait améliorer sensiblement la reconstruction. L'algorithme est celui du MAP (§2.2.3), soit jusqu'à convergence :

1. estimation $x^{(t+1)}$ de x par maximum de la loi *a posteriori* $p(x|\theta^{(t)}, y, z^{(t)}, s)$,
2. estimation $z^{(t+1)}$ par fusion de $z^{(0)}$ et $z'^{(t+1)}$ extrait de $x^{(t+1)}$.
3. estimation des $\widehat{\theta}^{(t+1)}$ par (2.14).

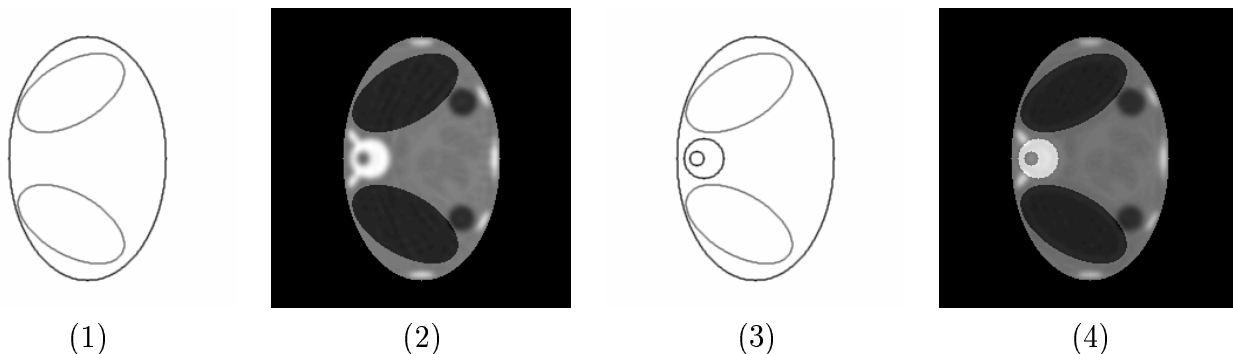


FIG. 2.7 – Contours connus (1), reconstruction (2), contours connus plus complets (3), reconstruction associée (4)

Une connaissance parfaite des contours permettrait une reconstruction de grande qualité (FIG.2.8). Mais cette connaissance *a priori* n'est pas réaliste (car transformerait notre problème mal-posé en la détermination des quelques valeurs du champ au sein des zones homogènes parfaitement définies).

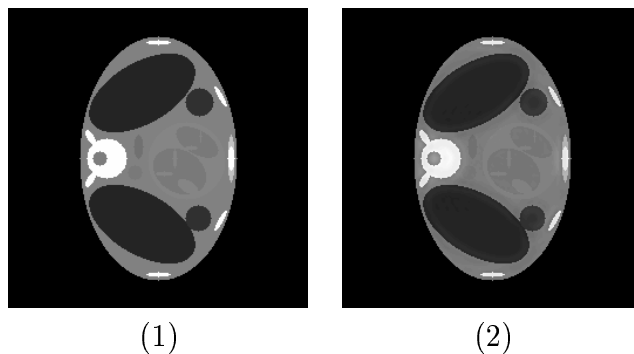


FIG. 2.8 – Image à reconstruire (1), reconstruction connaissant *tous* les contours (2)

L'algorithme défini plus haut permet d'obtenir des résultats encourageants (FIG.2.9). Les régions homogènes se séparent nettement, et les faibles variations au centre du fantôme commencent à se distinguer. Cette amélioration de la connaissance des contours permet donc une reconstruction *globale* de meilleure qualité.

Cependant, cette méthode ne fait qu'affirmer les contours existants, et n'est applicable qu'à partir de l'*a priori* supposant l'image à reconstruire constituée de zones homogènes. Elle peut être certainement améliorée par l'utilisation d'une méthode d'extraction plus performante que celle utilisée en simulations⁷.

⁷filtrage de Canny seuillé utilisé comme masque sur un filtrage linéaire, le tout rééchelonné à partir des contours donnés initialement.

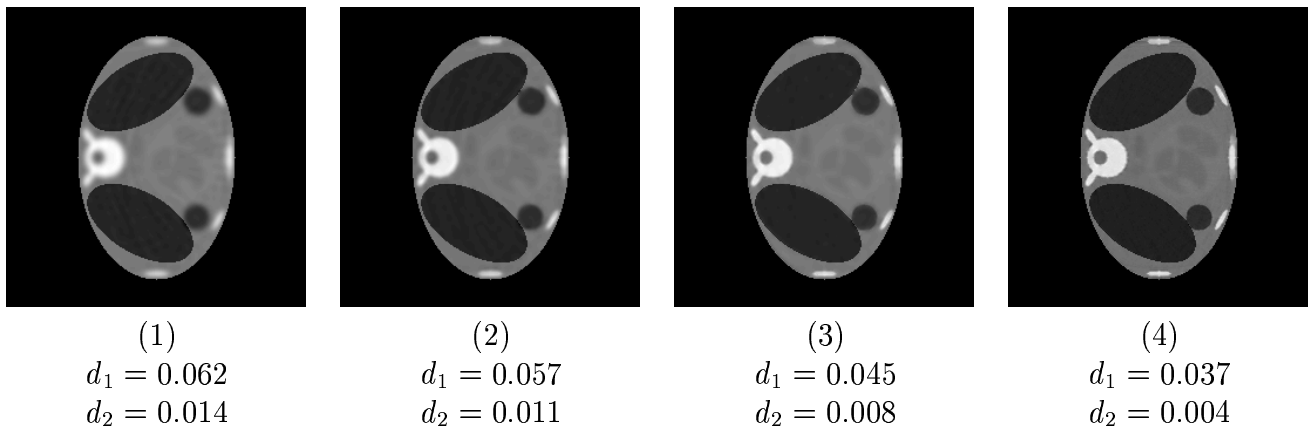


FIG. 2.9 – Reconstruction à partir des contours initiaux (1), reconstructions intermédiaires (2-3), reconstruction finale (4)

Conclusion

Le modèle de fusion homogène présenté écrit les trois jeux de données (tomographiques, échographiques et régions connues) comme ajout sur les opérations de mesures de trois bruits gaussiens. L'importance de chacun de ces bruits, donnée par leurs variances respectives, est supposée inconnue *a priori*.

Une connaissance approchée de celles-ci est nécessaire à la reconstruction [1]. Il a en effet été montré ici que leurs valeurs déterminent l'importance des sous-critères dans la reconstruction par maximum *a posteriori*, et l'amplitude des tirages des pixels dans la reconstruction par méthode de Monte-Carlo.

La conception d'un système automatique qui les estimerait, avec la reconstruction, est difficile par maximum *a posteriori*, une fois le bruit, partie intégrante du modèle et réalité pratique, ajouté aux mesures.

L'application de l'échantillonneur de Gibbs écrite ici permet une reconstruction d'une extrême robustesse au bruit, ceci pour un coût de calculs comparable à l'algorithme utilisé dans la maximisation *a posteriori*. La convergence, rapide, est obtenue sans valeur *a priori* sur l'image (simple rétroprojection) ou les variances (obtenues initialement à partir de celle-ci).

Les estimées de ces dernières permettent une évaluation, biaisée mais proche, du rapport signal sur bruit des données présentées au système. Le modèle pourrait ainsi être utilisé par un système expert (estimation des variances du bruit des signaux), ou comme simple algorithme de reconstruction.

Le modèle de fusion homogène présenté permet une écriture quadratique de la logvraisemblance de l'image à reconstruire, et ainsi l'utilisation directe de l'algorithme de Gibbs. L'introduction de lois non gaussiennes, comme l'accentuation du caractère homogène de la reconstruction par des *a priori* en exponentielles généralisées, pourrait être intéressante, mais au prix d'un échantillonnage par algorithme hybride (Gibbs-Hasting-Metropolis).

Enfin, l'utilisation de contours non binaires semble permettre une reconstruction de meilleure qualité. La procédure d'estimation devient néanmoins instable.

Annexe A

Champs de Gibbs-Markov

Soit E un ensemble (typiquement les indices $\{(k, l) \in \llbracket 1, M \rrbracket \times \llbracket 1, N \rrbracket\}$), $S = \mathcal{P}(E)$ ses parties et X une variable aléatoire (typiquement une image) définie sur E .

Système de voisinage

On définit un *voisinage* $d(s)$ de $s \in E$ comme un sous-ensemble de points ne contenant pas s :

$$d(s) = \{t \in E \mid t \neq s\} \quad (\text{A.1})$$

et un *système de voisinage* \mathcal{D} de E par la définition d'un voisinage pour chaque point, avec contrainte de réflexivité :

$$\mathcal{D} = \{d(s) \mid \forall t \in E, s \in d(t) \Leftrightarrow t \in d(s)\} \quad (\text{A.2})$$

On note \sim la *relation de voisinage* :

$$t \sim s \Leftrightarrow t \in d(s) \left(\Leftrightarrow s \in d(t) \right) \quad (\text{A.3})$$

Chaque élément s de E se voit définir l'ensemble de ses voisins $d(s)$, où s n'est pas voisin de lui-même, et t est voisin de s ssi s est voisin de t . Comme exemples, les boules de rayon n , $d(s) = \{t \in E \text{ avec } t \neq s, \|t - s\| < n\}$, en particulier (FIG.A.1) :

- $d(s) = \emptyset$: chaque point est indépendant de ses voisins
- 4-connexité (boule de rayon 2 pour la norme 1)
- 8-connexité (boule de rayon 2 pour la norme ∞)
- boule carrée de rayon 3 (boule de rayon 3 pour la norme ∞)
- $d(s) = E \setminus \{s\}$: tous les points sont voisins entre eux

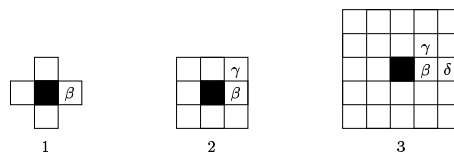


FIG. A.1 – Systèmes de voisinage : 4-connexité (1), 8-connexité (2) et boule carrée de rayon 3 (3)

Champs de Markov

Soit \sim une relation de voisinage définie sur E . X est un *champ de Markov pour la relation de voisinage \sim ssi* :

$$\forall s \in E, (x_s | x_{t \sim s}) \Rightarrow (x_s | x_{E \setminus \{s\}}) \quad (\text{A.4})$$

Soit une réalisation x_s du champ markovien X en un point s . Les réalisations de X hors du voisinage de s n'apportent donc pas d'information supplémentaire une fois connues celles au sein de celui-ci.

Cliques

Soit \sim une relation de voisinage définie sur E . On définit une *clique c pour la relation \sim* comme un ensemble d'éléments de E voisins 2 à 2 *pour le système de voisinage choisi* :

$$c \in S \quad / \quad \forall (\alpha, \beta) \in c^2, \alpha = \beta \quad \text{ou} \quad \alpha \sim \beta \quad (\text{A.5})$$

et un *système de cliques \mathcal{C} pour la relation \sim* comme l'ensemble des cliques c pour \sim . Le choix préalable de la relation de voisinage est *indispensable* à une définition correcte d'un système de cliques (FIG.A.2).

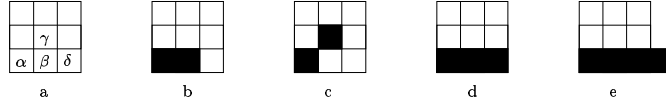


FIG. A.2 – (a) : $\alpha \sim \beta$ pour les relations de voisinage 1,2 et 3 (définis à la FIG.A.1), $\alpha \sim \gamma$ pour 2 et 3, $\alpha \sim \delta$ pour 3, (b) : clique pour 1,2 et 3 (c) : clique pour 2 et 3 (d) : clique pour 3 (e) : n'est pas une clique pour 1,2 ou 3

Champs de Gibbs

Soient \sim une relation de voisinage définie sur E et \mathcal{C} un système de cliques pour \sim . X est un *champ de Gibbs ssi* sa ddp est de la forme :

$$p(x) \propto \exp \left(- \Phi(x) \right) \quad \text{avec l'énergie} \quad \Phi(x) = \sum_{s \in S} \phi_s(x) \quad (\text{A.6})$$

et est un *champ de Gibbs de voisinage pour \mathcal{C} ssi* la somme des *potentiels ϕ_c* se restreint aux cliques :

$$\Phi(x) = \sum_{c \in \mathcal{C}} \phi_c(x) \quad (\text{A.7})$$

Théorème de Hammersley-Clifford

Soient \sim une relation de voisinage sur E et \mathcal{C} un système de cliques pour \sim . On a l'équivalence :

$$\boxed{X \text{ est de Gibbs de voisinage pour } \mathcal{C}}$$



$$\boxed{X \text{ est de Markov pour } \sim \text{ de densité strictement positive}}$$

L'hypothèse de *markovianité* que nous allons faire sur certain champs (images) est donc équivalente à assigner à ce champ une loi de Gibbs. Le choix des potentiels dépendra de l'*a priori* sur ces images (inhomogénéité, douceur, ...).

La modélisation markovienne de la loi *a priori* de l'image à reconstruire aide à la décision, car apporte une information supplémentaire à son sujet (FIG.2.1) (même si les images que l'on cherche à reconstruire sont très peu probables par simple génération aléatoire selon une loi de Markov). L'information de corrélation ajoutée est très générale. Des modèles plus complexes peuvent exister, mais pour une perte de généralité, une complexité accrue des calculs, et une amélioration peu significative des résultats (pour le modèle de fusion utilisé ici).

Exemples

Le caractère homogène d'une image peut être affirmé en assignant à la loi $p(x)$ une distribution de Gibbs-Markov d'énergie [3] :

$$U(x) = -\alpha \sum_s \left(\sum_{t \sim s} x_s x_t \right) \quad \text{avec } \alpha > 0 \quad (\text{A.8})$$

Pour le système de voisinage 4-connexité :

$$U(x) = -\alpha \sum_{k,l} x_{k,l} [x_{k,l-1} + x_{k,l+1} + x_{k-1,l} + x_{k+1,l}] \quad (\text{A.9})$$

et pour la boule carrée de rayon 3 (§...) :

$$U(x) = \alpha \sum_{k,l} x_{k,l} [\begin{aligned} &x_{k-2,l} \\ &+ 2x_{k-1,l-1} - 8x_{k-1,l} + 2x_{k-1,l+1} \\ &+ x_{k,l-2} - 8x_{k,l-1} + 20x_{k,l} - 8x_{k,l+1} + x_{k,l+2} \\ &+ 2x_{k+1,l-1} - 8x_{k+1,l} + 2x_{k+1,l+1} \\ &+ x_{k+2,l} \end{aligned}] \quad (\text{A.10})$$

Annexe B

Gradients matriciels

Le calcul de l'optimum du critère $J(x)$ (ie l'image reconstruite \hat{x} qui minimise $J(x)$) passe par le calcul de son gradient. La forme partiellement quadratique de $J(x)$ permet le calcul rapide de $\nabla J(x)$ en dérivant directement par rapport au vecteur image x , ceci à l'aide des *gradients matriciels*.

Définition

On définit la *dérivée de la quantité matricielle* A de taille (m, n) par rapport à la variable *matricielle* B de taille (p, q) par la matrice (mp, nq) , structurée en $m.n$ blocs $\partial A/\partial B_{ij}$ de dimension (p, q) , notée $\partial A/\partial B$ comme :

$$\left(\frac{\partial A}{\partial B}\right) = \left(\frac{\partial A}{\partial B_{ij}}\right)_{(i,j) \in \llbracket 1,p \rrbracket \times \llbracket 1,q \rrbracket} \quad (\text{B.1})$$

Ainsi, si x est un vecteur $(n,1)$ et $A(x)$ $(m,1)$ (A fonction $\mathbb{R}^n \rightarrow \mathbb{R}^m$), la matrice jacobienne de A s'écrit :

$$\boxed{J_A(x) = \frac{\partial A^\top}{\partial x}} \quad (\text{B.2})$$

Dans le cas particulier où $A : \mathbb{R}^n \rightarrow \mathbb{R}$:

$$\boxed{\nabla_A(x) = \frac{\partial A}{\partial x}} \quad (\text{B.3})$$

Pour A et B quelconques, on a :

$$\frac{\partial(A^\top)}{\partial(B^\top)} = \left(\frac{\partial A}{\partial B}\right)^\top \quad (\text{B.4})$$

Enfin, quelques propriétés de dérivées de fonctions linéaires, quadratiques et de composées vont permettre un calcul rapide du gradient $\nabla J(x)$ ((2.18)).

Propriétés

On a les trois propriétés suivantes, avec A matrice quelconque constante de x et x vecteur colonne :

$$\boxed{\begin{aligned} \frac{\partial}{\partial x}(x^\top A) &= \frac{\partial}{\partial(x^\top)}(Ax) = A \\ \frac{\partial}{\partial x}(x^\top Ax) &= Ax + A^\top x \\ \frac{\partial}{\partial x}((Ax)^\top(Ax)) &= 2A^\top Ax \end{aligned}} \quad (\text{B.5})$$

Gradient matriciel de composée

$$\text{Si : } \begin{cases} G : \mathbb{R}^n \rightarrow \mathbb{R}^m \\ x \mapsto G(x) \\ F : \mathbb{R}^m \rightarrow \mathbb{R} \\ y \mapsto F(y) \\ F \circ G : \mathbb{R}^n \rightarrow \mathbb{R} \\ x \mapsto (F \circ G)(x) \end{cases} \quad (\text{B.6})$$

alors :

$$\nabla(F \circ G)(x) = J_G(x) \cdot \nabla F(G(x)) \quad (\text{B.7})$$

Ainsi, en notation matricielle, avec x , F et G de tailles respectives $(n, 1)$, $(m, 1)$ et (m, n) :

$$\frac{\partial(F \circ G)}{\partial x} = \frac{\partial G^\top}{\partial x} \cdot \frac{\partial F}{\partial y}(G(x)) \quad (\text{B.8})$$

Applications

Vont être calculés ici les gradients $\partial/\partial x((y - Hx)^\top R^{-1}(y - Hx))$ et $\partial/\partial x(\Phi(Dx))$.

Calcul de $\nabla_x(y - Hx)^\top R^{-1}(y - Hx)$

$(y - Hx)^\top R^{-1}(y - Hx)$ apparaît lors de l'expression de la logvraisemblance d'une loi gaussienne de variable y , de moyenne Hx et de matrice de covariance R symétrique (§2.2.2).

$$\begin{aligned} \frac{\partial}{\partial x} \left((y - Hx)^\top R^{-1}(y - Hx) \right) &= \frac{\partial}{\partial z} \left(z^\top R^{-1}z \right) \quad \text{avec } z = y - Hx \\ &= \frac{\partial}{\partial x} (y - Hx)^\top \cdot \frac{\partial}{\partial z} \left(z^\top R^{-1}z \right) \\ &= -2H^\top \cdot \left(R^{-1}(y - Hx) \right) \end{aligned} \quad (\text{B.9})$$

En particulier,

$$\frac{\partial}{\partial x} \left(\|y - Hx\|^2 \right) = -2H^\top (y - Hx) \quad (\text{B.10})$$

Calcul de $\nabla_{\Phi}(Dx)$

$\Phi(Dx)$ est l'énergie associée à la loi markovienne sur l'image à reconstruire connaissant les contours. D est l'opérateur *symétrique* de dérivation.

$$\begin{aligned} \frac{\partial}{\partial x} \left(\Phi(Dx) \right) &= \frac{\partial}{\partial x} (Dx)^{\top} \cdot \frac{\partial \Phi}{\partial y} (Dx) \\ &= D^{\top} \cdot \frac{\partial \Phi}{\partial y} (Dx) \end{aligned} \quad (\text{B.11})$$

soit :

$$\nabla \left(\Phi(Dx) \right) (x) = D \cdot \nabla_{\Phi} (Dx) \quad (\text{B.12})$$

En particulier, si $\Phi = \|\cdot\|_2^2$, $\nabla_{\Phi}(y) = 2y$, soit :

$$\nabla \left(\Phi(Dx) \right) (x) = 2D \cdot Dx \quad (\text{B.13})$$

Le gradient de $\|Dx\|^2$ est donc la dérivée seconde (au sens du traitement numérique des images) de x .

Annexe C

Estimation de la variance d'une loi gaussienne

Modèle

Soit un champ x (de dimension N_x), lié à ses mesures y (N_y) par la relation $y = Hx + \epsilon$, avec $\epsilon \sim \mathcal{N}(0, R_\epsilon = 1/\sigma_\epsilon^2 I_{N_y})$. On pose $\theta = 1/\sigma_\epsilon^2$, hyperparamètre à estimer. On suppose à priori que :

$$\begin{cases} \theta \sim \gamma(a, b) \\ x \sim \mathcal{N}(m_x, R_x = \sigma_x^2 I_{N_x}) \end{cases} \quad (\text{C.1})$$

En effet, la vraisemblance $p(y|x, \theta)$ est gaussienne en x , gamma en θ . Il est donc naturel de choisir les lois *a priori* associées de mêmes formes. La loi gaussienne pour x traduit une variation possible autour d'une valeur moyenne connue *a priori*, et la loi inverse gamma pour σ_ϵ^2 interdit une variance négative, rend peu probable une grande variance, et favorise une valeur moyenne.

Ces connaissances *a priori* vont permettre d'estimer x et θ à partir des mesures y , par MAP.

Estimation de θ par maximum de $\theta|y$

A partir de la loi de Bayes, $p(\theta|y) \propto p(y|\theta).p(\theta)$, avec $y = Hx + \epsilon(\theta)$, et en notant que x (champ réel à reconstruire) et ϵ (bruit de modèle et mesures) sont indépendants :

$$y|\theta = Hx + \epsilon(\theta) \sim \mathcal{N}(Hm_x, HR_xH^\top + R_\epsilon) \quad (\text{C.2})$$

soit :

$$p(\theta|y) \propto \exp\left(-\frac{1}{2}(y - Hm_x)^\top (HR_xH^\top + R_\epsilon)^{-1}(y - Hm_x)\right).p(\theta) \quad (\text{C.3})$$

L'expression de $p(\theta|y)$ fait intervenir le calcul de l'inverse :

$$\theta(\lambda^2 HH^\top + I_{N_y})^{-1}, \text{ avec } \lambda = \sigma_x/\sigma_\epsilon \quad (\text{C.4})$$

Cette opération est très difficile, et le lemme d'inversion apporte peu.

L'estimation préalable de la variance du bruit σ_ϵ^2 à partir des données seules y est difficile

Ceci, dans le cas d'un *a priori* gaussien sur x . Dans un cas plus général, pour x markovien, intervient alors le calcul intégral :

$$p(\theta|y) \propto \int_x p(\theta, x|y) \propto \left[\int_x p(y|\theta, x).p(x) \right].p(\theta) \quad (\text{C.5})$$

pour un résultat similaire. La matrice (C.4) varie, à λ fixé, en θ , résultat qui se précise en ajoutant la connaissance de x .

Estimation de θ par maximum de $\theta|y, x$

En effet, $p(\theta|y, x) \propto p(y|\theta, x).p(\theta|x)$, soit, avec $(y|\theta, x) \sim \mathcal{N}(Hx, \sigma_\epsilon^2 I_{N_y})$:

$$\begin{aligned} p(\theta|y, x) &\propto \theta^{N_y/2} \exp \left[-\frac{\|y-Hx\|^2}{2\sigma_\epsilon^2} \right].p(\theta) \\ &\propto \theta^{N_y/2} \exp \left[-\frac{\|y-Hx\|^2}{2} \theta \right].\theta^{a-1} \exp(-b\theta) \\ &\propto \theta^{a+N_y/2-1} \exp \left[-\left(b + \frac{\|y-Hx\|^2}{2} \right) \theta \right] \end{aligned} \quad (\text{C.6})$$

L'hyperparamètre θ suit donc aussi *a posteriori* une loi gamma, $\theta|y, x \sim \gamma(a' = a + N_y/2, b' = b + \|y - Hx\|^2/2)$. Soit, par MAP :

$$\begin{aligned} \hat{\theta} &= \arg \max_{\theta} p(\theta|y, x) \\ &= (a' - 1)/b' \end{aligned} \quad (\text{C.7})$$

Estimation de x par maximum de $x|y, \theta$

En supposant x markovien d'énergie $U(x)$:

$$\begin{aligned} p(x|y, \theta) &\propto p(y|x, \theta).p(x|\theta) \\ &\propto p(y|x, \theta).p(x) \\ &\propto \exp \left(-\frac{\|y-Hx\|^2}{2\sigma_\epsilon^2} - U(x) \right) \\ &\propto \exp \left(-J(x) \right) \end{aligned} \quad (\text{C.8})$$

Soit, obtenu par un algorithme itératif de type gradient :

$$\begin{aligned} \hat{x} &= \arg \max_x p(x|y, \theta) \\ &= \arg \min_x J(x) \\ &= \arg \min_x \left(\theta \|y - Hx\|^2 + U(x) \right) \end{aligned} \quad (\text{C.9})$$

x et θ seront donc mis à jour successivement ((1.36)) :

$$\begin{cases} \hat{\theta} = (a + N_y/2 - 1)/(b + \|y - Hx\|^2/2) \\ \hat{x} = \arg \min_x J(x) \end{cases} \quad (\text{C.10})$$

C'est cette méthode qui va-t-être appliquée à la fusion homogène en reconstruction tomographique, dans le cas de l'estimation par MAP. Nous voyons ici dans l'écriture du critère $J(x)$ l'importance d'une valeur correcte de θ , poids du sous-critère associé. A rappeler que cet algorithme n'est pas équivalent à une maximisation jointe, et peut ne pas converger vers l'optimum global (F).

Annexe D

Méthodes de Monte-Carlo

Définition

Soit x une variable aléatoire de densité f . On appelle algorithme de MCMC (Monte-Carlo Markov Chains) toute méthode produisant, à partir d'un échantillon initial $y^{(0)}$, une chaîne de Markov $(y^{(t)})$ ergodique de loi stationnaire la distribution d'intérêt f [4].

Ainsi, si x est un champ à estimer, f sa loi *a posteriori*, les échantillons $(y^{(t)})$ construits seront (asymptotiquement) distribués selon cette même loi. La moyenne arithmétique des derniers sera alors un estimateur au sens de la MMSE de x .

Algorithme de Hasting-Metropolis

Lorsque la génération directe d'échantillons de loi f est possible, la suite :

$$y^{(t)} \sim f \tag{D.1}$$

sera bien sûr distribuée selon f . Bien souvent, la connaissance de la loi est limitée à une simple évaluation en chaque point, et un tirage selon f ne peut être directement réalisé. Est alors introduite une *loi instrumentale* g , pour donner l'algorithme de *Hasting-Metropolis* :

$$\left\{ \begin{array}{l} 1. \text{ générer } \tilde{y}^{(t)} \sim g \\ 2. \text{ calculer } \rho(y^{(t)}, \tilde{y}^{(t)}) = \min \left(\frac{f(\tilde{y}^{(t)}) g(y^{(t)})}{f(y^{(t)}) g(\tilde{y}^{(t)})}, 1 \right) \\ 3. \text{ choisir } y^{(t+1)} = \begin{cases} \tilde{y}^{(t)} & \text{avec la probabilité } \rho(y^{(t)}, \tilde{y}^{(t)}) \\ y^{(t)} & \text{sinon} \end{cases} \end{array} \right. \tag{D.2}$$

On retrouve la méthode d'échantillonnage précédente dans le cas où la loi instrumentale g s'identifie à la distribution d'intérêt f . Le taux d'acceptation des échantillons $\rho(y^{(t)}, \tilde{y}^{(t)})$ est dans ce cas toujours égal à 1. On remarque que la connaissance de la loi f à *une constante multiplicative près* suffit ((2.9)(2.13)).

La convergence de cette méthode est très lente, et sa rapidité dépendra du choix de la loi instrumentale. Plus proche elle sera de f , plus efficace sera l'algorithme.

Cas multidimensionnel

On notera (x_i) les composantes du champ x . La méthode de Hasting-Metropolis ne peut être utilisée telle qu'elle, par manque de performance, dans le cas multidimensionnel (comme des

images, champs de dimensions 256×256 dans les simulations), où chaque tirage nécessiterait l'évaluation de la loi f .

Echantillonneur de Gibbs

On se place dans le cas d'un champ multidimensionnel $x = (x_i)$ de loi $f = (f_i)$, d'échantillons directement accessibles. L'*échantillonneur de Gibbs* est une amélioration apportée à la méthode (D.1). La génération de se fait plus selon la loi conjointe, mais suivant les conditionnelles, en tenant compte des tirages précédents :

$$\left\{ \begin{array}{l} 1. \ y_1^{(t+1)} \sim f_1(y_1|y_2^{(t)}, \dots, y_n^{(t)}) \\ \text{i.} \ y_i^{(t+1)} \sim f_i(y_i|y_1^{(t+1)}, \dots, y_{i-1}^{(t+1)}, y_{i+1}^{(t)}, \dots, y_n^{(t)}) \\ \text{n.} \ y_n^{(t+1)} \sim f_n(y_n|y_1^{(t+1)}, \dots, y_{n-1}^{(t+1)}) \end{array} \right. \quad (\text{D.3})$$

Dans le cas où seules les évaluations de f sont disponibles (toujours connues à une constante multiplicative près), cette méthode ne peut être directement appliquée. Une méthode *hybride* entre Gibbs et Hasting-Metropolis permet de créer efficacement des échantillons d'une loi multidimensionnelle connue en simple évaluation.

Algorithme MCMC hybride

Il s'écrit, en notant $g = (g_i)$ la loi instrumentale :

$$\left\{ \begin{array}{l} 1. \text{ générer } \tilde{y}_i^{(t)} \sim g_i(y_i|y_1^{(t+1)}, \dots, y_{i-1}^{(t+1)}, y_{i+1}^{(t)}, \dots, y_n^{(t)}) \\ 2. \text{ calculer } \rho(y_i^{(t)}, \tilde{y}_i^{(t)}) = \min \left(\frac{f_i(\tilde{y}_i^{(t)}) g_i(y_i^{(t)})}{f_i(y_i^{(t)}) g_i(\tilde{y}_i^{(t)})}, 1 \right) \\ 2. \text{ choisir } y_i^{(t+1)} = \begin{cases} \tilde{y}_i^{(t)} & \text{avec la probabilité } \rho(y_i^{(t)}, \tilde{y}_i^{(t)}) \\ y_i^{(t)} & \text{sinon} \end{cases} \end{array} \right. \quad (\text{D.4})$$

(y_i) sera alors distribué asymptotiquement selon f .

Annexe E

Symétrie des opérateurs de dérivation

Les opérateurs $M = N^T N$ et H_{2i} , définis dans le modèle de fusion (§2.2.1), sont symétriques. Leur caractère de voisinage permet de les appliquer, en pratique, par une simple convolution. Le calcul du gradient du critère à minimiser (§2.2.3) s'effectue alors simplement.

Symétrie de la matrice de voisinage

$x_{i,j}$ [x_j en notation vectorielle] sera nul pour des indices (i, j) [j] hors dimensions, ceci afin de garder une écriture générique, malgré les effets de bords.

On remarque que Mx définie en (2.6) et (A.10) s'écrit $N.Nx$, avec $(Nx)_{i,j} = (1/4)(4x_{i,j} - x_{i,j-1} - x_{i,j+1} - x_{i-1,j} - x_{i+1,j})_{i,j}$, soit, en notations vectorielles, $(Nx)_j = (1/4)(4x_j - x_{j-M} - x_{j+M} - x_{j-1} - x_{j+1})_j$, pour les pixels respectifs central, gauche, droite, supérieur et inférieur (notations du §1.3.2). N est donc la matrice symétrique :

$$N = \frac{1}{4} \begin{pmatrix} 4 & -1 & 0 & \dots & 0 & -1 & 0 & \dots & 0 \\ -1 & \ddots & \ddots & & \ddots & \ddots & \ddots & & \vdots \\ 0 & \ddots & & & & & & & \\ \vdots & & & & & & & & \\ 0 & & & & & & & & \\ -1 & \ddots & & & & & & & \\ 0 & \ddots & & & & & & & \\ \vdots & & & & & & & & \\ 0 & \dots & & & & & & & \end{pmatrix} \quad (\text{E.1})$$

M est donc aussi symétrique.

Symétrie des opérateurs de dérivation

L'étude échographique a été décomposée en quatre opérateurs directionnels d'extraction de contours (§1.3.2). On a choisi ici :

$$\begin{cases} (H_{21}x)_j = \frac{1}{2}(2x_j - x_{j-1} - x_{j+1})_j \\ (H_{22}x)_j = \frac{1}{2}(2x_j - x_{j-M} - x_{j+M})_j \\ (H_{24}x)_j = \frac{1}{2}(2x_j - x_{j-M-1} - x_{j+M+1})_j \\ (H_{23}x)_j = \frac{1}{2}(2x_j - x_{j-M+1} - x_{j+M-1})_j \end{cases} \quad (\text{E.2})$$

Les quatre matrices de dérivation discrète H_{2i} sont donc, tout comme M , symétriques.

Equivalence matrices-masques

La taille $M \times N$ des images utilisées (§1.3.2) interdit l'utilisation directe des opérateurs M et H_{2i} (de tailles $MN \times MN$). Cependant :

Multiplier une image (sous forme vectorielle) par un opérateur matriciel



Filtrer une image (sous forme matricielle) par un masque de convolution

Soit, sous Matlab, dans le cas d'un opérateur matriciel F , de masque associé f :

$$\text{reshape}(F*x, M, N) = \text{conv2}(f, \text{reshape}(x, M, N), 'same')$$
 (E.3)

On peut ainsi définir les masques correspondant aux opérateurs N , M et H_{2i} par :

$$n = \frac{1}{4} \begin{pmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{pmatrix}$$
 (E.4)

$$m = \frac{1}{16} \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 2 & -8 & 2 & 0 \\ 1 & -8 & 20 & -8 & 1 \\ 0 & 2 & -8 & 2 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$
 (E.5)

$$h_{21} = \frac{1}{2} \begin{pmatrix} 0 & -1 & 0 \\ 0 & 2 & 0 \\ 0 & -1 & 0 \end{pmatrix}$$
 (E.6)

$$h_{22} = \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 \\ -1 & 2 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$
 (E.7)

$$h_{23} = \frac{1}{2} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$
 (E.8)

$$h_{24} = \frac{1}{2} \begin{pmatrix} 0 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 0 \end{pmatrix}$$
 (E.9)

Ainsi, lors du calcul du gradient du critère, les calcul Mx et $H_{2i}(H_{2i}x - z_i)$ pourront être réalisés très simplement par application de ces masques de convolution.

Enfin, on remarque que (2.5) :

$$h_2 = \sum_i h_{2i} = \frac{1}{8} \begin{pmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{pmatrix} \quad (\text{E.10})$$

est un masque de convolution pour une extraction omnidirectionnelle de contours. Le choix de paramètres θ_{2i} égaux limite l'introduction de données échographiques à une simple dérivation discrète.

Annexe F

JMAP hiérarchique

On cherche ici à estimer $z = (x, \theta)$ par maximum de sa loi *a posteriori* $p(z|y)$ (§1.5.2). La maximisation jointe (1.33) n'est pas possible en pratique, ni ses équivalents (1.34) et (1.35), par intégration de la loi $p(x, \theta|y)$. C'est la méthode par MAP hiérarchique (1.36) qui est utilisée dans les simulations. Cette dernière permet d'atteindre un maximum de la loi $p(x, \theta|y)$, qui peut être local seulement dans le cas d'une loi jointe non concave. Le résultat dépend alors de l'initialisation. En effet, en choisissant comme loi jointe (FIG.F.1) :

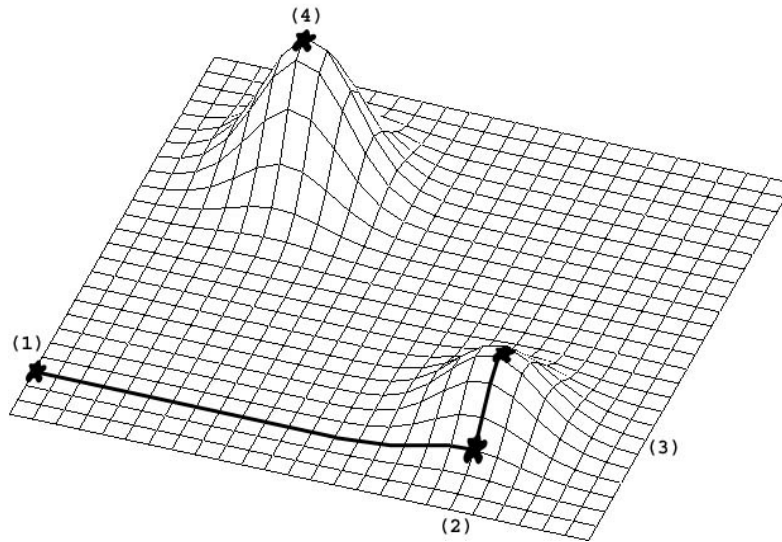


FIG. F.1 – Loi *a posteriori* $p(x, \theta|y)$

et comme valeurs initiales (1), une première maximisation selon $\theta|y, x$ mène au point de crête (2), une deuxième selon $x|y, \theta$ aboutit à l'extremum local (3). Les opérations suivantes ne délogeront pas l'estimée, pour lui faire atteindre le maximum global (4). Ainsi :

La méthode hiérarchique de maximisation *a posteriori* n'est pas nécessairement globale

Bibliographie

- [1] Ali Mohammad-Djafari. Hierarchical markov modeling for fusion of x ray radiographic data and anatomical data in computed tomography. In *Int. Symposium on Biomedical Imaging (ISBI 2002)*, 7-10 Jul., Washington DC, USA, July 2002.
- [2] Gabor T. Herman. *Image reconstruction from projections. The fundamentals of computerized tomography*. Academic Press, New York, NY, 1980.
- [3] Gerhard Winkler. *Image Analysis, Random Fields and Dynamic Monte Carlo Methods*. Springer Verlag, Berlin,, 1995.
- [4] Christian Robert. *Méthodes de Monté Carlo par Chaînes de Markov*. Economica, 1996.
- [5] Jérôme Idier, editor. *Approche bayésienne pour les problèmes inverses*. Traité IC2, Série traitement du signal et de l'image, Hermès, Paris, 2001.
- [6] S. R. Deans. *The Radon transform and some of its applications*. Wiley Interscience, New York, 1983.
- [7] Jérôme Idier. Convex half-quadratic criteria and interacting auxiliary variables for image restoration. Technical report, GPI-L2S, 1999.
- [8] Philippe Ciuciu and Jérôme Idier. A half-quadratic block-coordinate descent method for spectral estimation. Technical report, GPI-L2S, 2000.
- [9] Donald Geman and Chengda Yang. Nonlinear image recovery with half-quadratic regularization. 4(7) :932–946, July 1995.
- [10] Institut de Protection et de Sûreté Nucléaire. L'exposition accidentelle aux rayonnements ionisants.