Université de PARIS-SUD

MEMOIRE

Présenté pour obtenir le

DIPLOME D'HABILITATION A DIRIGER DES RECHERCHES

Entropie multi-échelle: définition et applications

par

Jean-Luc STARCK

Soutenue le 29 novembre 1999 devant le jury composé de:

G. Demoment	L2S- CNRS-Supélec	Président
P. Flandrin	CNRS-ENS Lyon	Rapporteur
F. Gamboa	Faculté d'Amiens	Examinateur
A. Mohammad-Djafari	L2S- CNRS-Supélec	Rapporteur
J.-C. Pesquet	L2S - CNRS - Supélec	Examinateur
S. Roques	Observatoire Midi-Pyrénées	Rapporteur
L. Vigroux	CEA-Service d'Astrophysique	Examinateur

Remerciements

Le travail que je présente ici est issu de nombreuses collaborations et discussions, et il ne m'est pas possible de citer tous ceux qui m'ont aidé au cours de ces dernières années. Qu'ils soient tous assurés de mon amitié.

Je voudrais tout de même remercier en particulier Pierre Borgeaud, Laurent Vigroux, et Catherine Cesarsky, pour m'avoir fait confiance et donné une place dans l'équipe ISOCAM. Avoir participer aux heures intenses du lancement d'ISO et à l'arrivée des premières données a été une grande joie. Je remercie par la même occasion toute l'équipe ISO, les collègues de l'IAS et de Vilspa, pour ces moments partagés.

Je remercie également mes collègues Ivan Valtchanov, Eric Pantin, Philippe Querre, et René Gastaud pour les nombreuses discussions animées concernant le traitement du signal, ainsi que Pierre-François pour sa précieuse aide concernant Linux, même s'il a une fâcheuse tendance à remplir mon disque dur.

Je tiens à remercier les membres du jury, Guy Demoment qui a accepté la présidence, Sylvie Roques, Ali Mohammad-Djafari, et Patrick Flandrin qui ont bien voulu être examinateur de mon habilitation, ainsi que Laurent Vigroux, Jean-Christophe Pesquet, et Fabrice Gamboa.

Mes remerciements vont aussi à Fionn Murtagh et à Albert Bijaoui pour notre collaboration qui a aboutit sur la concrétisation du projet *Multiresolution*.

Table des matières

Chapitre 1

Introduction

La motivation première de tous les travaux de recherche concernant le traitement des images réside dans la compréhension du contenu de ces dernières. En effet, si la manière dont l'information est distribuée dans nos données est modélisée correctement, alors les outils mathématiques de nos modèles vont permettre d'optimiser la représentation du signal, et d'améliorer tous les traitements, que ce soit à des fins de compression de données, de détection, de filtrage, ou de déconvolution. Parmi les outils apparus ces dernières années, on notera l'analyse fractale, les champs de Markov et la transformée en ondelettes. Les ondelettes sont particulièrement bien adaptées à l'analyse de structures de tailles variables, souvent superposées, et offrent donc un excellent cadre de travail pour l'imagerie astronomique. Étant donné que la totalité de mon travail de chercheur s'est déroulée dans des laboratoires d'astronomie, c'est autour des problèmes liés à l'astronomie que mes efforts se sont focalisés. Il faut toutefois noter que l'on retrouve souvent ces mêmes problèmes dans d'autres domaines (imagerie médicale, télé-détection, ...).

Le terme "entropie" est due à Clausius (1865), et le concept d'entropie a été introduit par Boltzmann en mécanique statistique dans le but de mesurer le nombre de possibilités microscopiques pour qu'un état macroscopique soit réalisé. Shannon (1948) a posé les fondements de la théorie de la communication en suggérant que l'information apportée par une mesure dépende de son nombre de possibilités d'apparition. Il a aussi montré que l'entropie peut être utilisée pour maximiser le transfert de bit sous une contrainte de qualité. L'entropie de Shannon nous permet de connaître exactement le nombre de bits minimum pour stocker un signal sans perte d'information, ce qui est évidemment fondamental dès que l'on veut stocker ou transmettre des données.

Jaynes (1957) a proposé de l'utiliser pour la déconvolution d'image interférométrique, dans le but de sélectionner, dans un ensemble possible de solutions, celle qui contient le moins d'information, ou suivant sa définition de l'entropie, celle dont l'entropie est maximale. En principe, la solution vérifiant une telle condition doit être la plus fiable. Cette méthode, appelée MEM (*Maximum Entropy Method*), a connu un grand succès, et beaucoup de travaux ont été menés durant ces 30 dernières années sur l'utilisation de l'entropie pour le problème général du filtrage de données ou de déconvolution (Bontekoe et al., 1994; Burg, 1978; Frieden, 1978a; Gull and Skilling, 1991; Narayan and Nityananda, 1986; Pantin and Starck, 1996; Skilling, 1989; Weir, 1992).

Nous résumons au second chapitre une partie de ces travaux, et nous introduisons dans un cadre ondelette une nouvelle forme de mesure d'information, l'entropie multiéchelle. Celle-ci a certaines propriétés qui nous semblent essentielles pour une bonne restauration. L'utilisation de la transformée en ondelettes permet en particulier de tenir compte de la corrélation des pixels. L'information mesurée devient alors une quantité qui dépend aussi de la disposition des pixels, et pas seulement de leurs valeurs.

Le troisième chapitre présente les applications de l'entropie multiéchelle. Nous verrons d'abord comment elle conduit à un bon compromis entre seuillage doux et seuillage dur dans le cas du filtrage, puis comment elle permet de pallier les difficultés que présente MEM pour la déconvolution. Nous montrons finalement comment elle peut également servir à l'analyse de fluctuation du fond du ciel, proposant ainsi une alternative à des outils comme la fonction de corrélation à deux points.

Enfin, j'ai également participé de manière active aux efforts de réflexion et de développement de méthodes pour répondre aux besoins du Service d'Astrophysique (SAP) du CEA, principalement sur le projet ISO (*Infrared Space Observatory*), satellite lancé en novembre 1995, où le service d'astrophysique du CEA était maître d'oeuvre de l'un des quatre instruments, la caméra ISOCAM. Les programmes cosmologiques (*deep survey, Hubble Deep Field*, ...) ont nécessité en particulier de gros efforts pour parvenir à détecter de manière automatique les objets les plus faibles. Le quatrième chapitre synthétise l'ensemble de ces travaux, en soulignant bien les parties liées à un encadrement d'un thésard ou d'un stagiaire. En conclusion, une sélection d'articles publiés est présentée dans le dernier chapitre.

Chapitre 2

Principe de l'Entropie Multi-Echelle

2.1 Entropie et restauration d'images

2.1.1 Introduction

Dans les sciences physiques, on modélise en général un ensemble de données Y par le mélange d'un signal utile et d'un bruit, qui est supposé suivre dans la plupart des cas une loi de distribution Gaussienne, Poissonnienne, ou encore une combinaison des deux.

Si on note X ($X = \{X_1, ..X_N\}$, N étant le nombre de pixels), les données que l'on aurait reçues en l'absence de bruit, la relation entre Y et X en présence d'un bruit additif est:

$$Y = X + B \tag{2.1}$$

Chercher X à partir de Y est un problème de filtrage.

Si la réponse de l'instrument de mesure est invariante spatiallement, la relation objet-image est:

$$Y = P * X + B \tag{2.2}$$

où P est la réponse instrumentale, que l'on appelle aussi PSF (*Point Spread Function*), et X représente cette fois le signal original. Chercher X connaissant P et Y est un problème de déconvolution.

Le théorème de Bayes permet d'établir la relation de densité de probabilité conditionnelle suivante:

$$\mathrm{p}(X|Y) = \frac{\mathrm{p}(Y|X).\mathrm{p}(X)}{\mathrm{p}(Y)} \tag{2.3}$$

Pour calculer le maximum a posteriori (MAP) de l'objet, on doit maximiser

$$\underset{X}{\mathrm{Max}}\ \mathrm{p}(X|Y) = \underset{X}{\mathrm{Max}}\ \frac{\mathrm{p}(Y|X).\mathrm{p}(X)}{\mathrm{p}(Y)} \tag{2.4}$$

$\mathrm{p}(Y|X)$ est la densité de probabilité conditionnelle d'obtenir les données connaissant le signal original X, (c'est en fait la distribution du bruit). Le dénominateur est indépendant de X et est considéré comme une constante. Les deux termes importants sont donc $\mathrm{p}(Y|X)$ et $\mathrm{p}(X)$. Le premier est dérivé d'un modèle du bruit. Par exemple, dans le cas d'un bruit blanc Gaussien d'écart type σ et de moyenne nulle, on a:

$$\mathrm{p}(Y|X) \propto \exp(-\sum_{k=1}^{N} \frac{(Y_k - (P * X)_k)^2}{2\sigma^2}) \tag{2.5}$$

p(X) est la loi de probabilité a priori de la solution X. En l'absence d'information sur la solution X, à l'exception de sa positivité, elle peut être dérivée de l'entropie par (Shannon, 1948; Jaynes, 1957; Shore and Johnson, 1981; Johnson and Shore, 1984; Bijaoui, 1984; Mohammad-Djafari, 1994; Mohammad-Djafari, 1998)

$$\mathrm{p}(X) = \exp(-\alpha H(X)) \tag{2.6}$$

où $H(X)$ est l'entropie de X (on décrira dans la suite la manière de la calculer).

Compte tenu des données, la solution la plus probable est obtenue en maximisant $\mathrm{p}(X|Y)$. En prenant le logarithme de l'équation 2.3, on doit maximiser

$$\ln(\mathrm{p}(X|Y)) = -\alpha H(X) + \ln(\mathrm{p}(Y|X)) - \ln(\mathrm{p}(Y)) \tag{2.7}$$

La solution du maximum a posteriori est indépendante du dernier terme, celui-ci étant constant et pouvant donc être supprimé. Dans le cas d'un bruit blanc Gaussien, la solution est déterminée en minimisant

$$J(X) = \sum_{k=1}^{N} \frac{(Y_k - (P * X)_k)^2}{2\sigma^2} + \alpha H(X) = \frac{\chi^2}{2} + \alpha H(X) \tag{2.8}$$

c'est une combinaison linéaire de deux termes: l'entropie du signal, et une quantité correspondant au χ^2, mesurant l'écart entre les données et le modèle prédit. α est un paramètre qui peut être vu comme un paramètre de Lagrange ou une valeur fixant le poids relatif entre la fidélité aux données et l'entropie H. Une valeur élevée de α impose plus de poids sur le terme lié à l'entropie par rapport au χ^2, et la solution est plus lisse. Le paramètre α fixe donc le degré de *douceur* sur la solution.

2.1.2 Le concept d'entropie

Shannon (Shannon, 1948) a défini, dans le cadre de la théorie de la communication, l'entropie d'une image par

$$H_s(X) = -\sum_{k=1}^{N_b} p_k \log p_k \tag{2.9}$$

où $X = \{X_1, ..X_N\}$ est une image contenant des valeurs entières, N_b est le nombre de valeurs possibles que peut prendre un pixel X_k (par exemple 256 pour une image codée sur huit bits), et les p_k sont dérivés de l'histogramme de X:

$$p_k = \frac{\#X_j = k}{N} \tag{2.10}$$

$\#X_j = k$ donnant le nombre de pixels $X_j = k$. Cette forme d'entropie est particulièrement pertinante car elle permet de connaître le nombre minimum de bits nécessaire au codage de l'image sans perte d'information. Elle est donc extrèmement utilisée dans toutes les applications de compression ou de transfert de données.

Si l'image contient des valeurs réelles, on peut construire un histogramme des valeurs X_i, en utilisant un intervalle adéquat Δx, en comptant le nombre m_k de fois que chaque intervalle $(x_k, x_k + \Delta x)$ apparaît lors des N occurrences. Alors la probabilité qu'une valeur appartienne à un intervalle k est $p_k = \frac{m_k}{N}$, et chaque valeur des données a une probabilité p_k.

L'entropie est minimum et égale à zéro quand le signal est constant, et augmente en présence de fluctuations. Utiliser cette expression de l'entropie dans l'équation 2.8 conduit à minimiser:

$$J(X) = \frac{\chi^2}{2} + \alpha H_s(X) \tag{2.11}$$

On a alors une méthode de restauration par minimum d'entropie.

L'inconvénient de cette approche est le suivant: à cause du nombre fini d'occurrences, l'estimation de la valeur p_k est erronée par une quantité proportionnelle à $m_k^{-\frac{1}{2}}$ (Frieden, 1978b). Cette erreur devient significative quand m_k est petit. De plus, cette forme d'entropie n'est pas facile à utiliser pour la restauration de données, car le gradient de l'équation 2.8 est difficile à calculer.

Nous avons décrit ci-dessus l'entropie de Shannon, car elle est à la base de tous les travaux sur l'entropie, même si à ma connaissance, elle n'a jamais été utilisée pour de la restauration d'images, sans doute pour les raisons citées auparavant. Cependant, une fonctionnelle d'entropie peut être construite de différentes manières (voir (Aczel and Daroczy, 1978)), et nous ne citons dans la suite que les plus répandues.

Frieden (1978a) a proposé d'assimiler la valeur des pixels à une densité de probabilité p_k, et l'entropie devient alors:

$$H_f(X) = -\sum_{k=1}^{N} X_k \ln(X_k) \tag{2.12}$$

Deux autres formes d'entropie, souvent citées dans la littérature, sont l'entropie de Burg (1978) $H_b(X) = -\sum_{k=1}^{N} \ln(X_k)$, et surtout celle de Gull et Skilling (1991):

$$H_g(X) = \sum_{k=1}^{N} X_k - M_k - X_k \ln(\frac{X_k}{M_k}) \tag{2.13}$$

M étant un modèle donné. Cette définition de l'entropie a l'avantage d'être nulle et maximale quand X est égal au modèle M (le modèle choisi est généralement un signal ou une image constante).

Chacune de ces entropies peut être utilisée, et elle correspond à différentes distributions de probabilité que l'on peut associer à une image (Narayan and Nityananda, 1986). (voir (Frieden, 1978a; Skilling, 1989) pour plus de détails).

Toutes ces mesures d'entropie sont négatives (si $X_k > 1$), et maximum quand le signal est constant. Elles sont négatives car un terme additif n'est pas pris en compte, ce qui n'a aucune importance pour la minimisation de la fonctionnelle. La recherche de la solution la plus lisse implique alors une une maximisation de l'entropie, au lieu d'une minimisation, et l'équation à minimiser dans le cas d'un bruit Gaussien est:

$$J(X) = \frac{\chi^2}{2} - \alpha H(X) \tag{2.14}$$

Le signe devant le second terme a été artificiellement inversé (voir équation 2.8), ce qui est naturel si on veut que la meilleure solution soit la plus lisse. Ceci revient tout de même à redéfinir la loi de probabilité de X (voir équation 2.6):

$$\mathrm{p}(X) = \exp(\alpha H(X)) \tag{2.15}$$

Cette inversion de signe fait que ces trois formes d'entropie (H_f, H_b, et H_g) conduisent à des méthodes de restauration par maximum d'entropie, appelées MEM (*Maximum Entropy Method*).

Narayan et Nityanda (1986) ont comparé plusieurs fonctions d'entropie, et ont conclu finalement qu'elles sont toutes valables si elles ont les bonnes propriétés, c'est à dire qu'elles forcent la positivité, et qu'elles ont une dérivée seconde négative qui empêche les oscillations. Ils ont aussi montré que les résultats varient fortement avec le niveau de fond, et que les structures négatives (c'est à dire sous le niveau de fond, zone absorption pour une image, bande d'absorption dans un spectre, etc.) sont relativement mal restaurées, de même que les structures compactes.

L'entropie de Gull et Skilling pose le problème de l'estimation d'un modèle M. Dans le cas des images astronomiques, un bon choix pour M est le fond du ciel, mais son estimation n'est pas forcément aisée. Il est montré (Bontekoe et al., 1994) que la solution dépend du choix de M.

La détermination du paramètre α n'est pas non plus une tâche facile, et est en réalité un problème sérieux que pose la méthode du maximum d'entropie. Dans le fameux algorithme MAXENT de Skilling et Gull, le choix de α est conditionné par le fait qu'il doit satisfaire la contrainte $\chi^2 = N$ quand la déconvolution est terminée, N étant le nombre de degrés de liberté du système (le nombre de pixels dans le problème de déconvolution d'image). Mais ce choix conduit systématiquement à une sous estimation des données (Titterington, 1985) qui est clairement apparente pour les problèmes d'imagerie avec une petite PSF. En réalité, le χ^2 est supposé varier dans l'intervalle $N \pm \sqrt{2N}$ d'une réalisation des données à une autre. Du point de vue de la quantification du Maximum d'Entropie (Skilling, 1989), la valeur optimale de α est déterminée en incluant sa loi de probabilité $\mathrm{p}(\alpha)$ dans l'équation de Bayes et en maximisant alors la probabilité d'avoir α, connaissant les données et le modèle m. En pratique, une valeur de α trop grande donne un résultat qui est trop régularisé avec une importante perte de résolution. Une valeur trop petite conduit à une solution mal régularisée, avec beaucoup d'artefacts. Prendre un modèle uniforme, pour l'image a priori, lisse les discontinuités, ce qui peut sembler inacceptable pour des images astronomiques, qui contiennent souvent des étoiles ou des objets très compacts. La méthode standard de maximum d'entropie ne semble pas être très appropriée pour cette sorte d'images qui contient en même temps des hautes et des basses fréquences.

Pour résoudre ces problèmes, tout en gardant le concept de maximum d'entropie, certaines améliorations ont été proposées. Après avoir remarqué que des pixels voisins dans l'image reconstruite pouvaient avoir des valeurs très différentes dans des zones supposées uniformes (Charter, 1989), Gull et Skilling ont introduit le concept de fonction cachée h (*hidden function*) et de fonction de corrélation intrinsèque (ICF pour *Intrinsic Correlation Function*) C (C étant une fonction Gaussienne ou une spline cubique) dans l'algorithme MAXENT.

L'ICF décrit une échelle minimum de corrélation dans la solution O qui est obtenue par

$$O = C * S \tag{2.16}$$

Ceci correspond à imposer une résolution minimale à la solution O. Puisque la solution cachée S n'est pas spatialement corrélée, elle peut être régularisée par l'entropie

$$H_g(S) = \sum_{k=1}^{N} S_k - M_k - S_k \ln(\frac{S_k}{M_k}) \tag{2.17}$$

Puisque beaucoup d'échelles sont présentes dans des images astronomiques, Weir a proposé la méthode MEM-multi-cannaux (*Multi-channel Maximum Entropy Method*) (Weir, 1991; Weir, 1992) qui consiste à utiliser un ensemble d'ICFs ayant différentes longueurs d'échelles, chacune définissant un canal. La solution de l'espace "visible" (par opposition à l'espace des solutions cachées) est maintenant formée par une somme pondérée d'images O_j de l'espace visible:

$$O = \sum_{j=1}^{N_c} p_j O_j \tag{2.18}$$

où N_c est le nombre de canaux. Comme précédemment, chaque solution O_j est supposée être le résultat d'une convolution entre une image cachée S_j avec un filtre passe-bas (ICF) C_j:

$$O_j = C_j * S_j \tag{2.19}$$

Mais une telle méthode a plusieurs inconvénients:

1. La solution dépend de la largeur des ICFs (Bontekoe et al., 1994).
2. Il n'y a pas de manière rigoureuse de fixer les poids p_j (Bontekoe et al., 1994).
3. Le temps de calcul augmente linéairement avec le nombre de pixels.
4. La solution obtenue dépend du choix des modèles m_j ($j = 1 \ldots N_c$) qui sont choisis indépendamment des canaux.

Bontekoe et al. (1994) ont introduit une pyramide de résolution (*Pyramid Maximum Entropie*). L'approche pyramidale permet d'avoir des ICFs de taille constante, et le temps de calcul est réduit. Il a été montré (Bontekoe et al., 1994) que tous les poids pouvaient être fixés à 1 ($p_j = 1$ pour chaque canal).

Cette méthode élimine les trois premiers inconvénients, et donne une meilleure reconstruction des structures piquées et des structures lisses. Mais par contre, aux deux derniers inconvénients, un nouveau s'est ajouté: comme les images O_j ont des tailles différentes (dû à l'approche pyramidale), la solution O est reconstruite par duplication des pixels des sous-images O_j de chaque canal. Ceci produit des artefacts dûs à la création de hautes fréquences qui sont incompatibles avec le spectre réel de la vrai image $\hat{O}$. Toutefois ce problème peut être en partie résolu en reconvoluant le canal par l'ICF, ou en utilisant une fonction d'interpolation pour la reconstruction. L'introduction de la pyramide de résolution a résolu un certain nombre de problèmes et apporté beaucoup d'améliorations à la méthode MEM, mais elle a aussi posé d'autres questions. Dans le but de dériver le modèle à partir des valeurs physiques, Pantin & Starck (1996) ont introduit la transformée en ondelettes, et écrit

$$H(O) = \frac{1}{\sigma_I^2} \sum_{j=1}^{l} \sum_{k=1}^{N_j} \sigma_j (w_{j,k} - M_{j,k} - |w_{j,k}| \ln \frac{|w_{j,k}|}{M_{j,k}}) \tag{2.20}$$

où σ_I est l'écart type du bruit dans l'image I, et σ_j est l'écart type du bruit à l'échelle j (σ_j est déduit de σ_I grâce aux propriétés de la transformée en ondelettes). L'entropie multi-échelle est la somme de l'entropie à différentes échelles.

Les coefficients w_j sont des coefficients d'ondelette, et on prend la valeur absolue de w_j dans cette définition car les w_j peuvent être positifs ou négatifs, et un signal négative contient aussi de l'information dans sa transformée en ondelettes.

L'avantage d'une telle définition de l'entropie est le fait que l'on puisse utiliser de précédents travaux concernant la transformée en ondelettes et la restauration d'images (Murtagh et al., 1995b; Starck and Murtagh, 1994; Starck and Bijaoui, 1994a). Le comportement du bruit a déjà été étudié dans l'espace ondelette et on peut estimer l'écart type du bruit σ_j à l'échelle j. Ces estimations peuvent alors être introduites dans les modèles $M_{j,k}$

$$M_{j,k} = k_m \sigma_j \tag{2.21}$$

Le modèle m_j à l'échelle j représente la valeur prise par un coefficient d'ondelette en l'absence de signal significatif et en pratique, il doit être petit devant une valeur significative. En prenant m_j comme étant une fraction du bruit σ_j, m_j peut être considérée comme une sorte de limite physique en dessous de laquelle un signal ne peut pas être distingué du bruit ($k_m = \frac{1}{100}$).

2.1.3 Conclusion

On a pu voir que beaucoup d'études ont été menées dans le but de choisir la fonctionnelle à minimiser. Mais la question qui mérite d'être posée est la suivante: "quelle fonction peut être considérée comme une bonne mesure de l'information (entropie) pour la restauration des données ?"

En essayant de répondre à cette question, nous considérerons tout d'abord qu'un signal X est une somme de plusieurs composantes:

$$X = S + F + B \tag{2.22}$$

où S est le signal utile, F le fond du ciel, et B le bruit.

La fonction entropie H de X devrait satisfaire ces critères (Starck et al., 1998b):

1. **L'information dans un signal uniforme est nulle ($S = 0$, $B = 0$ et $F = \mathrm{Cst}$).**

2. **La quantité d'information dans X est indépendante du niveau de fond de l'image ($H(X)$ indépendant de F).**

3. **La quantité d'information est dépendante du bruit ($H(X)$ dépendant de B). Un signal donné ne donne pas la même information selon que le bruit est petit ou grand.**

4. **L'entropie doit fonctionner de la même manière pour X ayant la valeur $F+\epsilon$ ou $F-\epsilon$. Etant donné que $H(X)$ doit être indépendant de F (point 2), ça revient à écrire que $H(X)$ doit dépendre de la valeur absolue de S et non de S.**

5. **La quantité d'information est dépendante de la corrélation des pixels dans S. Si S possède de grandes structures au-dessus du bruit, il contient beaucoup d'information. En générant aléatoirement un nouveau signal à partir des valeurs des pixels de S, ces structures disparaissent et le signal contient moins d'information (tout en gardant les mêmes valeurs des pixels).**

Ces cinq points ne sont pas des axiomes, mais plutôt des tendances que devrait suivre la fonctionnelle d'entropie afin d'effectuer une bonne restauration des données.

La figure 2.1 montre Saturne (à gauche), et une image obtenue en distribuant de manière aléatoire les valeurs des pixels de l'image de Saturne. Ces deux images ont donc exactement le même histogramme. Pour une personne qui n'est pas impliquée dans le traitement des images, la seconde image contient moins d'information que la première. En revanche, pour quelqu'un travaillant dans le domaine de la transmission de données, il est clair que la seconde image demandera plus de bits pour une transmission sans perte, et de son point de vue, la seconde image contient alors plus d'information. Les fonctions standard d'entropie présentées auparavant donnent exactement la même valeur pour les deux. Dans le cadre de la restauration de données, toutes les fluctuations liées au bruit sont sans intérêt, et ne contiennent pas réellement d'information. D'un point de vue physique, ceci est une bonne raison pour laquelle les définitions standards de l'entropie semblent mal adaptées à la mesure d'information pour la restauration de signaux.

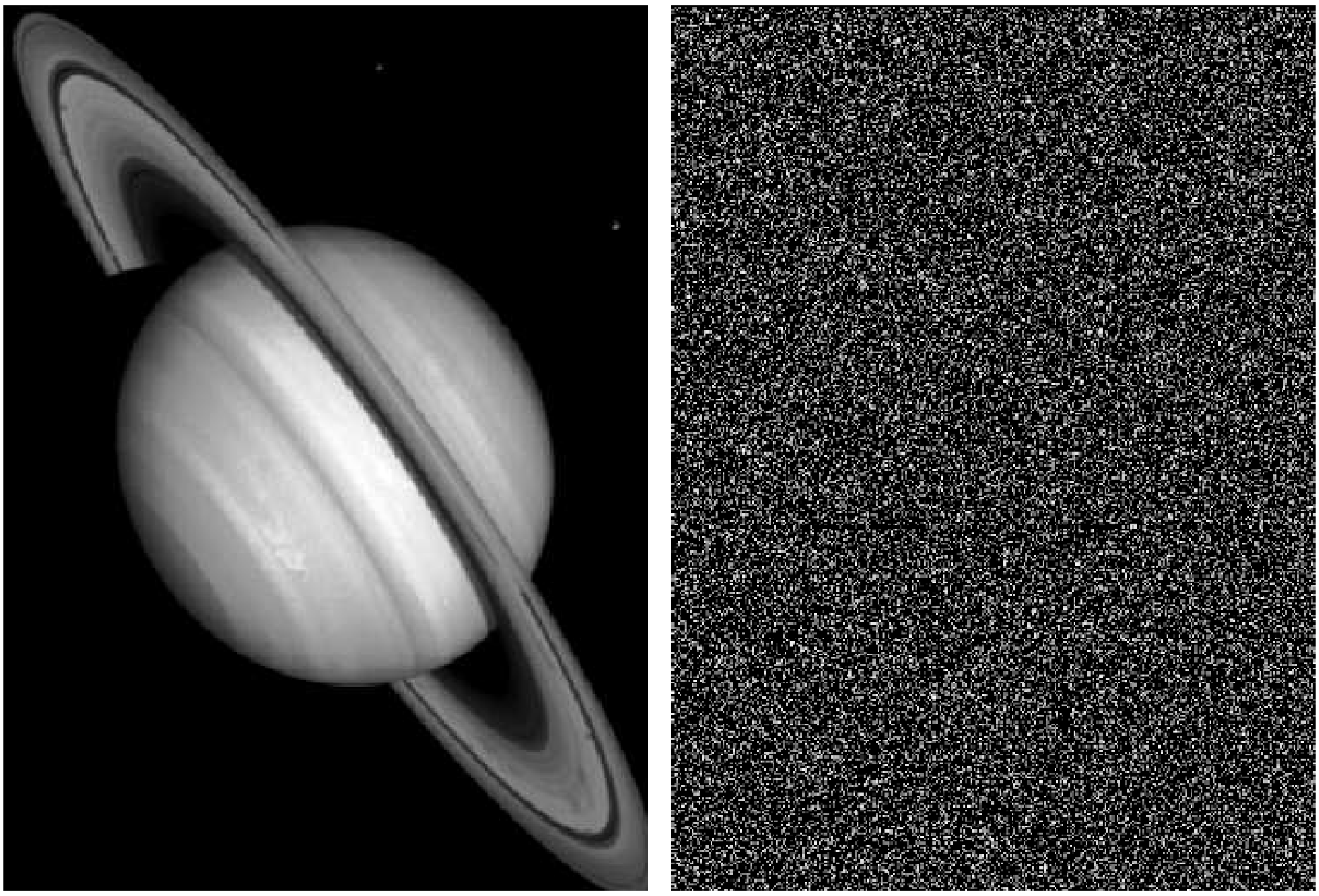

Figure 2.1: L'image de Saturne (gauche) et les mêmes données, mais distribuées aléatoirement (droite). Ces deux images ont la même entropie en utilisant les méthodes standards.

2.2 Entropie dépendant d'un modèle de bruit

2.2.1 Information et coefficients d'ondelettes

Pour ce qui concerne la restauration de signal, le bruit est le problème principal. Ceci signifie que nous ne devrions pas considérer la probabilité d'apparition d'une valeur dans l'image, mais plutôt sa probabilité d'être due à du signal (ou à du bruit). Si on considère une variable x qui suit une densité de probabilité $p(x)$, on peut définir l'information relative à x par $-\ln(p(x))$, et le signal X peut être considéré comme un ensemble de variables individuelles X_k (pixels), chacune suivant la même distribution. Alors l'information contenue dans les données peut être mesurée par $-\sum_{k=1}^{N}\ln(p(X_k))$. Si X_k suit une distribution Gaussienne de moyenne nulle, et d'écart type σ, on a

$$H(X) = \sum_{k=1}^{N} \frac{X_k^2}{2\sigma^2} + \mathrm{Cst} \tag{2.23}$$

L'énergie donne une bonne mesure de l'information (Starck et al., 1998b). Mais de nombreux critères ne sont pas rempli par une telle entropie (corrélation entre les pixels, indépendance par rapport au niveau de fond, etc.). Il semble difficile de dériver une bonne distribution de probabilité directement à partir des valeurs des pixels qui remplissent nos contraintes.

Il n'en est pas de même pour des données transformées. Ceci est particulièrement vrai si on utilise la transformée en ondelettes, car de nombreux travaux concernant le filtrage ont montré comment dériver des seuils de signification dans les différentes échelles à partir de densités de probabilité (Starck et al., 1995; Moulin and Liu, 1999; Donoho and Johnstone, 1993; Krim et al., 1992; Chipman et al., 1997; Amato and Vuza, 1997). Donc, une solution est d'introduire le concept de multirésolution dans l'entropie. On considère alors l'information contenue dans les données comme la somme d'information à différents niveaux de résolution j. La transformée en ondelettes W d'un signal X par un algorithme rapide contient un ensemble de coefficients d'ondelettes $w_{j,k}$ (j étant l'indice d'échelle), ainsi que des coefficients c_k représentant le signal à très basse résolution (voir (Mallat, 1998; Starck et al., 1998a) pour plus de détails sur la transformée en ondelettes). Si le nombre d'échelles utilisé est suffisamment grand, on peut supposer que les coefficients c_k portent de l'information essentiellement sur le fond de l'image, et non pas sur le signal utile. L'entropie de X ne doit donc être mesurée qu'à partir des coefficients d'ondelettes $w_{j,k}$.

Grâce aux propriétés de la transformée en ondelettes, les coefficients $w_{j,k}$ ont une moyenne nulle. A partir d'un modèle de bruit, on peut dériver la densité de probabilité d'un coefficient d'ondelettes, en supposant que sa valeur est due au bruit (et pas au signal). Si on utilise une transformée en ondelettes orthogonales, les coefficients en ondelettes sont indépendants, et l'entropie devient (Starck et al., 1998b)

$$H(X) = \sum_{j=1}^{l} \sum_{k=1}^{N_j} h(w_{j,k}) \tag{2.24}$$

avec $h(w_{j,k}) = -\ln p(w_{j,k})$. Pour un bruit Gaussien, on a

$$H(X) = \sum_{j=1}^{l} \sum_{k=1}^{N_j} \frac{w_{j,k}^2}{2\sigma_j^2} + \mathrm{Cst} \tag{2.25}$$

σ_j est le bruit à l'échelle j. On voit que l'information est proportionnelle à l'énergie des coefficients d'ondelettes. Si un coefficient d'ondelettes est grand par rapport au bruit, alors sa probabilité est très faible, et il est porteur de beaucoup d'information. On prendra la liberté dans la suite d'omettre la constante qui n'a pas d'effet dans le calcul des solutions dans les problèmes de restauration. On vérifie aisément que cette entropie vérifie toutes nos contraintes énoncées dans la section 2.1.2. Si on considère deux signaux S_1, S_2, dérivés d'un troisième S_0 par ajout d'un bruit:

$$\begin{aligned} S_1 &= S_0 + N_1(\sigma_1) \\ S_2 &= S_0 + N_2(\sigma_2) \end{aligned} \quad (2.26)$$

Alors on a:

$$\text{si } \sigma_1 < \sigma_2 \text{ alors } H(S_1) > H(S_2) \quad (2.27)$$

et une image uniforme a une entropie nulle.

Notre définition de l'entropie est complètement dépendante d'un modèle de bruit. Étant donné un signal S, et une hypothèse de bruit Gaussien d'écart type σ, on ne mesure pas la même information si on considère une autre valeur d'écart type de bruit, ou un autre modèle de bruit (Poissonnien par exemple).

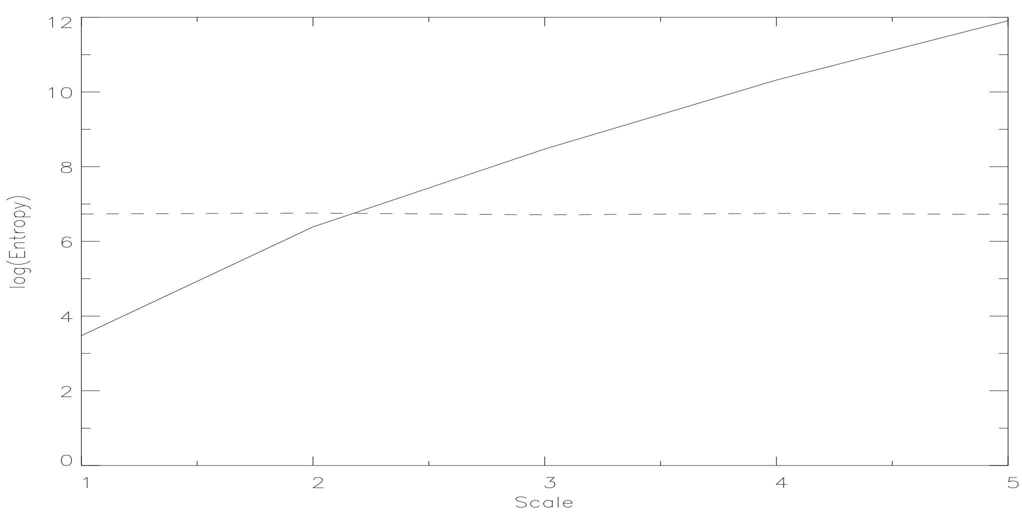

Figure 2.2: Entropie Multi-échelle de Saturne (trait continu), et entropie multi-échelle de l'image dont les pixels sont distribué aléatoirement (trait en pointillé).

La figure 2.2 montre l'information mesurée à chaque échelle pour l'image de Saturne et l'image générée aléatoirement à partir des valeurs des pixels de l'image de Saturne. L'entropie finale est la somme des entropies à chaque échelle. On voit que pour la seconde image (trait pointillé), l'information en fonction de l'échelle est une droite de pente nulle, tandis que pour l'image de Saturne, l'information augmente avec l'échelle. La transformée en ondelettes permet de tenir compte de la corrélation des pixels.

2.2.2 Information signal et information bruit

Dans la section précédente, nous avons vu comment il est possible de mesurer l'information liée à un coefficient d'ondelettes. En supposant que le signal X soit toujours composé des trois éléments S,F,B (X=$S+F+B$, signal utile, fond, et bruit), nous avons vu que H est bien indépendant du fond F. Notre mesure H d'information reste "polluée" par le bruit. Séparons maintenant notre information H en deux composantes, dont une (H_S) correspond à la partie qui ne peut pas être liée à du bruit, et l'autre (H_N) à celle certainement liée au bruit: on a

$$H(X) = H_S(X) + H_N(X) \tag{2.28}$$

On définit dans la suite H_S comme étant l'information signal, et l'autre H_N l'information bruit. Il doit être clair que le bruit ne contient pas d'information, et ce que l'on appelle information bruit est une quantité que l'on mesure comme information par l'entropie multi-échelle, mais qui n'a pas d'intérêt selon nous.

L'information h relative à un coefficient d'ondelettes $w_{j,k}$ est $-\ln(p(w_{j,k}))$. Si le coefficient d'ondelettes est petit, sa valeur peut être due à du bruit, et h doit affecter à H_N. Si le coefficient d'ondelettes est grand, comparé à l'écart type du bruit, h ne peut pas être due à du bruit, et h doit être affectée à H_S. h doit donc être distribué suivant H_N où H_S, en se basant par exemple sur la probabilité $P_n(w_{j,k})$ que le coefficient est dû à du bruit, ou la probabilité $P_s(w_{j,k})$ qu'il est dû à du signal. On a $P_s(w_{j,k}) = 1 - P_n(w_{j,k})$. Dans le cas Gaussien, on estime $P_n(w_{j,k})$ par (Starck et al., 1998b)

$$\begin{aligned} P_n(w_{j,k}) = \mathrm{Prob}(W > \mid w_{j,k} \mid) &= \frac{2}{\sqrt{2\pi}\sigma_j} \int_{|w_{j,k}|}^{+\infty} \exp(-W^2/2\sigma_j^2) dW \\ &= \mathrm{erfc}(\frac{\mid w_{j,k} \mid}{\sqrt{2}\sigma_j}) \end{aligned}$$

Pour tous les coefficients d'ondelettes $w_{j,k}$, on doit estimer les fractions h_n et h_s de h qui doivent être affectées à H_n et H_s. L'information signal et l'information bruit sont donc définies par

$$\begin{aligned} H_s(X) &= \sum_{j=1}^{l} \sum_{k=1}^{N_j} h_s(w_{j,k}) \qquad (2.29) \\ H_n(X) &= \sum_{j=1}^{l} \sum_{k=1}^{N_j} h_n(w_{j,k}) \end{aligned}$$

Notons que $H_s(X) + H_n(X)$ est toujours égal à $H(X)$.

N1-MSE

Une première approche pour dériver h_s et h_n de P_s et P_n est de considérer P_s et P_n comme des poids sur l'information h (Starck et al., 1998b). On a alors:

$$\begin{aligned} h_s(w_{j,k}) &= P_s(w_{j,k}) h(w_{j,k}) \qquad (2.30) \\ h_n(w_{j,k}) &= P_n(w_{j,k}) h(w_{j,k}) \end{aligned}$$

et les informations bruit et signal sont

$$\begin{aligned} H_s(X) &= \sum_{j=1}^{l} \sum_{k=1}^{N_j} h_s(w_{j,k}) \\ H_n(X) &= \sum_{j=1}^{l} \sum_{k=1}^{N_j} h_n(w_{j,k}) \end{aligned} \tag{2.31}$$

ce qui conduit pour le cas Gaussien:

$$\begin{aligned} H_s(X) &= \sum_{j=1}^{l} \sum_{k=1}^{N_j} \frac{w_{j,k}^2}{2\sigma_j^2} \mathrm{erf}(\frac{\mid w_{j,k} \mid}{\sqrt{2}\sigma_j}) \\ H_n(X) &= \sum_{j=1}^{l} \sum_{k=1}^{N_j} \frac{w_{j,k}^2}{2\sigma_j^2} \mathrm{erfc}(\frac{\mid w_{j,k} \mid}{\sqrt{2}\sigma_j}) \end{aligned} \tag{2.32}$$

On se référera à ces fonctions par le nom N1-MSE dans la suite.

N2-MSE

Avec la précédente définition de l'entropie, l'information relative aux très grands coefficients d'ondelettes est complètement affectée au signal. Ceci permet d'exclure les coefficients de grand rapport signal sur bruit (SNR), d'une régularisation lors d'un processus de restauration. Si on veut tenir compte de l'information de bruit, même sur les coefficients significatifs, l'information de bruit relatif à un coefficient d'ondelettes doit être définie différemment. L'idée que nous développons maintenant pour calculer h_s et h_n est la suivante: on imagine que l'information h est une somme de composants élémentaires d'information dh, chacun ayant une probabilité d'être de l'information signal ou de l'information bruit (Starck et al., 1998b). Pour comprendre le principe, considérons deux coefficients u et w ($w > u$) avec un bruit Gaussien ($\sigma = 1$). L'information relative à w est $h(w) = w^2$. En faisant varier u de 0 à w avec un pas du, l'information $h(u)$ augmente jusqu'à devenir égale à $h(w)$. Au fur et à mesure que u se rapproche de w, la probabilité que l'écart $w - u$ soit lié au bruit augmente, et donc l'incrément en information est de plus en plus contaminé par le bruit. En mettant un poids à l'incrément en information en fonction de la probabilité que $w - u$ soit lié à du bruit, on a

$$h_n(w_{j,k}) = \int_0^{\mid w_{j,k} \mid} P_n(\mid w_{j,k} \mid - u) (\frac{\partial h(x)}{\partial x})_{x=u} du \tag{2.33}$$

est l'information bruit relatif à $w_{j,k}$, et

$$h_s(w_{j,k}) = \int_0^{\mid w_{j,k} \mid} P_s(\mid w_{j,k} \mid - u) (\frac{\partial h(x)}{\partial x})_{x=u} du \tag{2.34}$$

est l'information signal de $w_{j,k}$.

Dans le cas Gaussien, on a

$$h_n(w_{j,k}) = \frac{1}{\sigma_j^2} \int_0^{\mid w_{j,k} \mid} u \ \mathrm{erfc}(\frac{\mid w_{j,k} \mid - u}{\sqrt{2}\sigma_j}) du \tag{2.35}$$

$$h_s(w_{j,k}) = \frac{1}{\sigma_j^2} \int_0^{\mid w_{j,k} \mid} u \ \mathrm{erf}(\frac{\mid w_{j,k} \mid - u}{\sqrt{2}\sigma_j}) du \tag{2.36}$$

et l'information de bruit et de signal sont

$$\begin{aligned} H_s(X) &= \sum_{j=1}^{l} \sum_{k=1}^{N_j} h_s(w_{j,k}) \\ H_n(X) &= \sum_{j=1}^{l} \sum_{k=1}^{N_j} h_n(w_{j,k}) \end{aligned} \quad (2.37)$$

Dans la suite, on appellera N2-MSE cette convention.

Les équations 2.32 et 2.30 conduisent à deux différentes manières de mesurer l'information dans les échelles. Suivant les applications, on choisira celle qui semble le mieux adaptée. La première implique que l'on utilise toute l'information relative à un fort coefficient d'ondelettes, permettant par exemple de ne pas perdre de flux dans la restauration d'images astronomiques. La seconde approche traite mieux le bruit (car on en tient compte même sur les forts coefficients), mais tolère une certaine perte d'information de l'ordre de grandeur du niveau de bruit.

LOG-MSE

L'information multi-échelle telle que nous l'avions défini dans (Pantin and Starck, 1996) (nous l'appellerons LOG-MSE dans la suite) peut être intégrée dans notre approche en définissant h par:

$$h(w_{j,k}) = \frac{\sigma_j}{\sigma_X^2}[w_{j,k} - M_j - \mid w_{j,k} \mid \log(\frac{\mid w_{j,k} \mid}{k_m \sigma_j})] \quad (2.38)$$

où σ_X est l'écart type du bruit dans les données. Et h_n est défini par:

$$h_n(w_{j,k}) = A(P_n(w_{j,k}))h(w_{j,k}) \quad (2.39)$$

où A est une fonction qui prend les valeurs 0 ou 1, en fonction de la valeur de $P_n(w_{j,k})$:

$$A(P_n(w_{j,k})) = \begin{cases} 1 & \text{si } P_n(w_{j,k}) > \epsilon \\ 0 & \text{si } P_n(w_{j,k}) \leq \epsilon \end{cases} \quad (2.40)$$

Les coefficients d'ondelettes significatifs imposent que $A(P_n(w_{j,k}))$ soit égale à 0, et ne contribuent pas à H_n. Cela signifie qu'utiliser H_n dans un processus de régularisation n'aura un effet qu'aux échelles et aux positions où aucun coefficient significatif n'aura été détecté.

2.2.3 Conclusion

En pratique, on préfère N1-MSE et N2-MSE pour plusieurs raisons: tout d'abord, le modèle utilisé dans l'équation 2.38 est un peu artificiel, et il y a une indétermination quand un coefficient d'ondelettes est égal à zéro. Ensuite, LOG-MSE semblent difficile à généraliser à d'autres classes de bruit, ce qui n'est pas le cas de N1-MSE et N2-MSE. N2-MSE a l'avantage d'estimer la part due au bruit dans l'information mesurée h, même sur les forts coefficients d'ondelettes.

En ce qui concerne les points cités en section 2.1.3:

1. *L'information dans un signal uniforme est nulle.*
   C'est vrai pour N1-MSE et N2-MSE. Dans le cas de LOG-MSE, c'est vrai en prenant $\sigma = 0$.

2. *La quantité d'information dans un signal est indépendante du niveau de fond de l'image.*
   C'est vrai avec les trois, car la dernière échelle (c'est à dire le plan lissé de basse résolution obtenu par la transformée en ondelettes) n'est jamais prise en compte dans le calcul de l'entropie.

3. *La quantité d'information est dépendante du bruit.*
   Quelque soit la méthode, ce sont des coefficients normalisés qui sont pris en compte, c'est donc toujours vrai.

4. *L'entropie doit fonctionner de la même manière pour un signal ayant la valeur $B + \epsilon$ ou $B - \epsilon$.*
   Pour N1-MSE et N2-MSE, c'est vrai par construction, car l'entropie est calculée à partir des valeurs absolues ou des carrés des coefficients d'ondelettes. Pour LOG-MSE, ce point n'est pas vérifié, car il y un terme $w_{j,k}$ qui apparaît. Mais ça n'a pas de conséquence sur la solution, car c'est la dérivée de l'entropie qui est utilisée dans les calculs et ce terme devient une constante.

5. *La quantité d'information est dépendante de la corrélation dans un signal.*
   Comme l'entropie est calculée à partir des coefficients d'ondelettes, ce point est toujours vérifié.

Les formes d'entropies proposées ici vérifient donc bien les points que nous considérons comme indispensables dans le cadre de la restauration des données.

Chapitre 3

Application de l'Entropie Multi-échelle

3.1 Filtrage

3.1.1 Introduction

La transformée en ondelettes (WT), en séparant les composantes d'un signal dans les différentes échelles, est un outil bien adapté à la suppression du bruit. Il y a beaucoup d'algorithmes de transformée en ondelettes bi-dimensionnelle (Starck et al., 1998a), et le plus connu est sans aucun doute la transformée orthogonale proposée par Mallat (1989), et son extension bi-orthogonale (Cohen et al., 1992). Ces méthodes sont basées sur le principe de la non-redondance de l'information dans les données transformées. D'autres algorithmes existent néanmoins, comme celui de Feauveau (1990) (qui est une transformée orthogonale, mais qui utilise une ondelette isotrope), ou l'algorithme à trous qui est non-orthogonal et fournit un ensemble de données très redondant (Holschneider et al., 1989). Toutes ces méthodes ont des avantages et des inconvénients. Suivant le contenu des données chacune de ces transformées peut être considérée comme mieux adaptée.

Une fois que la transformée est choisie, le second point est d'estimer le comportement du bruit dans les données transformées. Les transformées linéaires ont dans ce cas l'avantage de permettre une estimation robuste de l'écart type du bruit dans les échelles. Enfin, il faut procéder au traitement des coefficients d'ondelettes, et là encore, plusieurs stratégies peuvent être employées, comme le seuillage "doux" (*soft thresholding*) ou "dur" (*hard thresholding*) (Donoho and Johnstone, 1993; Donoho, 1993), et dans ce dernier cas, le niveau de seuillage (Starck and Bijaoui, 1994a; Moulin and Liu, 1999; Donoho and Johnstone, 1993; Krim et al., 1992; Chipman et al., 1997; Amato and Vuza, 1997; Nason, 1994; Nason, 1996; Nason, 1996).

Nous décrivons rapidement dans cette section les algorithmes qui peuvent être utilisés pour une décomposition en multirésolution (on parle ici de décomposition en multirésolution plutôt que de transformée en ondelettes, car on peut aussi obtenir des échelles de résolution par des algorithmes non linéaires qui ne sont pas des transformées en ondelettes). D'une manière générale, le choix d'un algorithme plutôt qu'un autre est en fait déjà l'introduction d'un modèle a priori des données. On peut donc parfaitement parler de modèle de vision pour une transformée en ondelettes ou un algorithme permettant une représentation multi-échelles de l'information. Pour des images de type astronomique ou médical, où les structures sont relativement diffuses et où la notion de contour ne correspond pas vraiment à grand chose, un algorithme isotrope est certainement bien adapté, tandis que pour des images comme un portrait, un algorithme privilégiant certaines directions permet souvent de bien restaurer les contours.

3.1.2 Multirésolution et filtrage

Cette section décrit les différentes stratégies possibles pour le filtrage des coefficients d'ondelettes.

Choix du modèle de vision

- **transformée en ondelettes (bi-) orthogonale.**
  Cette transformée en ondelettes (Mallat, 1989), souvent appelée *Fast Wavelet Transform* (FWT), est certainement la plus utilisée parmi les différents algorithmes de transformée en ondelettes. Elle permet une représentation non redondante de l'information. Une introduction à ce type de transformée peut être trouvée dans (Daubechies, 1992; Mallat, 1998).

  Un exemple est la transformée en ondelettes de Haar qui consiste à utiliser une ondelette définie par

  $$\begin{array}{ll} \psi(x) = 1 & \text{si } 0 \leq x < \frac{1}{2} \\ \psi(x) = -1 & \text{si } \frac{1}{2} \leq x < 1 \\ \psi(x) = 0 & \text{sinon} \end{array}$$

- **la transformée en ondelettes de Feauveau .**
  Feauveau (1990) a introduit une analyse en quinconce basée sur les travaux de Adelson (1987). Cette analyse est non dyadique et permet une décomposition de l'image avec un facteur de résolution égal à $\sqrt{2}$. Avec cette méthode, nous n'avons qu'une image ondelettes par échelle (tandis que nous en avons trois avec l'algorithme orthogonal et bi-orthogonal).

- **L'algorithme à trous.**
  La transformée en ondelettes d'une image c_0 par cet algorithme produit, à chaque échelle j, une image $\{w_j\}$. Celle-ci a le même nombre de pixels que l'image de départ c_0. c_0 peut être écrit comme la somme de toutes les échelles et du dernier plan lissé c_p

  $$c_0 = c_p + \sum_{j=1}^{p} w_j \tag{3.1}$$

  et un pixel à la position x, y peut être aussi écrit comme la somme des coefficients d'ondelettes à cette position, plus le coefficient du plan lissé:

  $$c_0(x,y) = c_p(x,y) + \sum_{j=1}^{p} w_j(x,y) \tag{3.2}$$

- **La transformée médiane multirésolution.**
  La transformée médiane est non linéaire, et offre l'avantage d'être un lissage robuste en présence de pixel ayant des valeurs extrêmes (comme par exemple les pixels touchés par des rayons cosmiques). La transformée médiane multirésolution (Starck et al., 1996e) consiste en une série de lissages par la médiane de l'image d'entrée avec des fenêtres de plus en plus

grande. La différence entre deux images lissées produit une nouvelle échelle de résolution. Celle-ci n'est par forcément de moyenne nulle, comme dans le cas de la transformée en ondelettes. Pour une image de valeur entière, cette transformée peut être faite en entier. Comme pour l'algorithme à trous, l'image originale peut être exprimée comme la simple somme des échelles et du dernier plan lissé.

Filtrage dans l'espace ondelette

Nous présentons maintenant les différentes approches pour traiter les coefficients multirésolutions, une fois que les données ont été transformées.

- **Seuillage dur**
  Il consiste à mettre à 0 tous les coefficients qui ont une valeur absolue plus petite qu'un seuil donné T_j ($T_j = k\sigma_j$, où j est l'échelle du coefficient d'ondelettes, σ_j l'écart type du bruit à l'échelle j, et k une constante généralement choisie égale à 3).

- **Seuillage doux**
  Le seuillage doux consiste à remplacer chaque coefficient d'ondelettes par la valeur $\tilde{w}$ où

$$\begin{aligned} \tilde{w}_j &= sgn(w_j)(\mid w_j \mid - T_j) \text{ si } \mid w_j \mid \geq T_j \\ &= 0 \text{ sinon} \end{aligned} \tag{3.3}$$

- **Niveau de seuil de Donoho**
  Donoho (Donoho and Johnstone, 1993; Donoho, 1993) a suggéré de prendre $T_j = \sqrt{2\log(n)}\sigma_j$ (où n est le nombre de pixels) au lieu de la valeur $k\sigma_j$. Ceci conduit à une nouvelle manière de faire de seuillage doux et dur.

- **Filtrage de Wiener Multiresolution**
  Le filtrage de Wiener multirésolution(Starck and Bijaoui, 1994a) consiste à multiplier tous les coefficients w_j d'une échelle j par

$$\alpha_j = \frac{S_j}{S_j + N_j} \tag{3.4}$$

  où S_j et N_j sont respectivement la variance du signal et du bruit à l'échelle j.

- **Filtrage de Wiener Hiérarchique**
  Le filtrage de Wiener hiérarchique (Starck and Bijaoui, 1994a) essaye d'introduire une prédiction dans l'estimation de $\tilde{w}_j$. Cette prédiction est obtenue à partir du coefficient w_h à la même position mais à l'échelle suivante.

$$\tilde{w}_j = \frac{H_j}{N_j + H_j + Q_j} w_j + \frac{N_j}{N_j + H_j + Q_j} w_h \tag{3.5}$$

  avec:

$$Q_j = \frac{H_j N_j}{S_j} \tag{3.6}$$

  où H_j est la variance de l'image obtenue en prenant la différence entre l'échelle j et la suivante $j + 1$.

- **Seuillage dur hiérarchique**
  Le seuil utilisé ici , T_h (Starck and Bijaoui, 1994a), est égal à $T_j = k\sigma_j$ si $\mid w_j \mid \geq T$, et $T_h = Tf(\mid \frac{w_h}{S_h} \mid)$ sinon (S_h est l'écart type de w_h). La fonction $f(a)$ doit retourner une valeur comprise entre 0 et 1. Une fonction possible pour f est:
  - $f(a) = 0$ si $a \geq k$
  - $f(a) = 1 - \frac{1}{k}a$ si $a < k$

Dans un seuillage doux, on considère que tous les coefficients sont bruités, et qu'ils doivent par conséquent tous être corrigés. La philosophie d'un seuillage dur est qu'il faut au contraire laisser intact les coefficients de fort rapport signal sur bruit (SNR) afin de ne perdre que le minimum d'information importante. Suivant les applications, on préfère l'une au l'autre approche. En effet, si seul l'aspect visuel compte, le seuillage doux est en général meilleur car il y a moins d'artefacts visuels. Dans le cas d'images astronomiques, un seuillage doux est catastrophique car il se traduit par une perte de photométrie des objets. Cette perte de flux est clairement visible quand on regarde les résidus (ou l'erreur), c'est à dire les données moins les données filtrées. En ce qui concerne le niveau de seuillage, le seuil de Donoho correspond à un risque mimimum. Plus la taille de l'échantillon augmente, plus le risque augmente, il est donc normal que le seuil dépende du nombre de pixels. Le seuil à $k\sigma$ correspond à une probabilité de fausse détection, c'est à dire de ne pas seuiller un coefficient qui est certainement lié à du bruit. Un seuil à 3σ correspond à 0.27 % de fausses détections. Le filtrage de Wiener Multirésolution est dérivé de l'hypothèse que le signal et le bruit suivent des lois gaussiennes. Cette hypothèse n'est en général pas vérifiée en ce qui concerne le signal. Le filtrage de Wiener hiérarchique introduit la notion de dépendance inter-échelle ou de persistance de l'information à travers les échelles. En effet, si à une échelle donnée, on trouve des coefficients de fort SNR, il est très probable que l'on en trouve aussi à l'echelle suivante. Cette persistance de l'information est actuellement utilisée par les méthodes de compression d'images les plus performentes (Shapiro, 1993; Said and Pearlman, 1996). Le seuillage dur hierarchique fait varier le niveau du seuil en fonction de cette dépendance inter-échelle. Ces méthodes seront évaluées dans la suite sur deux images différentes.

3.1.3 Principe du filtrage par entropie multiéchelle

Définition

Le problème du filtrage d'un ensemble de données D peut être exprimé sous cette forme: On cherche une solution $\tilde{D}$ telle que la différence entre D et $\tilde{D}$ minimise l'information due au signal, et $\tilde{D}$ minimise l'information due au bruit (Starck and Murtagh, 1999)

$$J(\tilde{D}) = H_s(D - \tilde{D}) + H_n(\tilde{D}) \tag{3.7}$$

De plus, le degré de douceur de la solution peut être controllé en ajoutant un paramètre:

$$J(\tilde{D}) = H_s(D - \tilde{D}) + \alpha H_n(\tilde{D}) \tag{3.8}$$

En faisant l'hypothèse que les coefficients d'ondelettes sont indépendants, on minimise en pratique (Chambolle et al., 1998) pour chaque coefficient d'ondelettes $w_{j,k}$:

$$j_c(\tilde{w}_{j,k}) = h_s(w_{j,k} - \tilde{w}_{j,k}) + \alpha h_n(\tilde{w}_{j,k}) \tag{3.9}$$

En utilisant une transformée en ondelettes orthogonales, le passage de l'équation 3.8 à l'équation 3.9 est rigoureux car l'hypothèse sur les coefficients d'ondelettes est vérifiée. Dans le cas d'une transformée redondante, l'hypothèse n'est évidemment pas vérifiée, mais l'expérience montre que l'incidence sur les résultats est faible, et l'utilisation d'une transformée redondante conduit en général à de meilleurs résultats.

$j_c(\tilde{w}_{j,k})$ peut être obtenu par une méthode quelconque de minimisation. Dans nos exemples, nous avons utilisé une simple dichotomie.

La figure 3.1 montre le résultat quand on minimise la fonctionnelle j_c avec différentes valeur de α, et un écart type de bruit égal à 1. Les coefficients d'ondelettes corrigés sont présentés en fonction des coefficients d'ondelettes. De la courbe du haut à celle du bas, α est respectivement égal à 0, 0.1, 0.5, 1, 2, 5, 10. Plus la valeur de α est grande, plus le coefficient d'ondelettes corrigé est réduit. Quand α est nul, il n'y a pas de régularisation et les données sont inchangées.

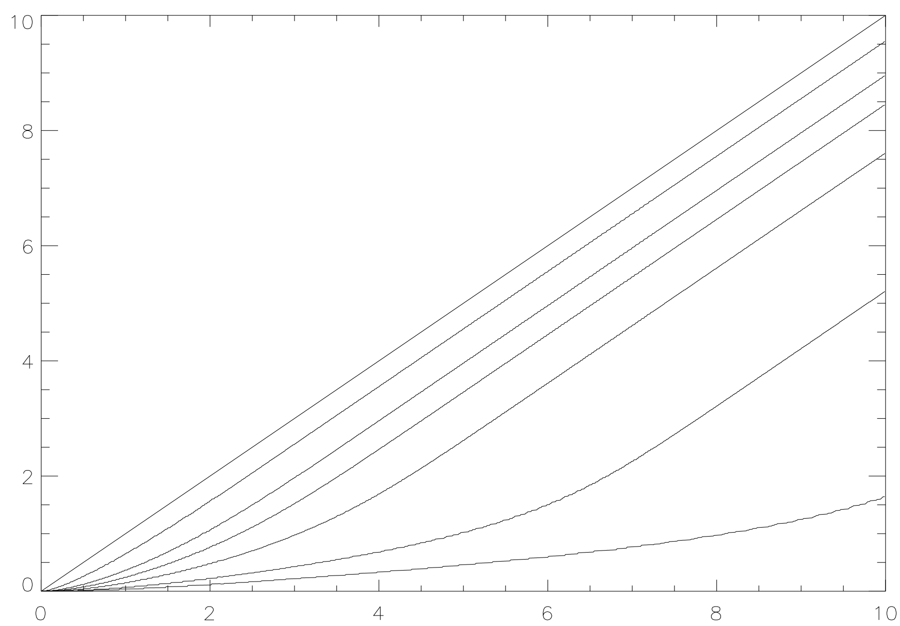

Figure 3.1: Coefficient d'ondelettes corrigé en fonction du coefficient d'ondelettes pour différentes valeurs de α (de la courbe du haut à celle du bas, α est respectivement égal à 0,0.1,0.5, 1, 2, 5,10).

Le paramètre de régularisation

Le paramètre α peut être utilisé de différentes manières (Starck and Murtagh, 1999):

- Il peut être fixé à une valeur donnée (paramètre utilisateur): $\alpha = \alpha_u$. Cette méthode conduit à filtrage très rapide en utilisant les optimisations présentées ultérieurement.
- Il peut être calculé sous la contrainte que le résidu vérifie certaines caractéristiques particulières. Par exemple, dans le cas d'un bruit Gaussien, on attend un résidu dont l'écart type est compatible avec l'estimation de l'écart type du bruit. Dans ce cas, $\alpha = \alpha_c \alpha_u$. Le

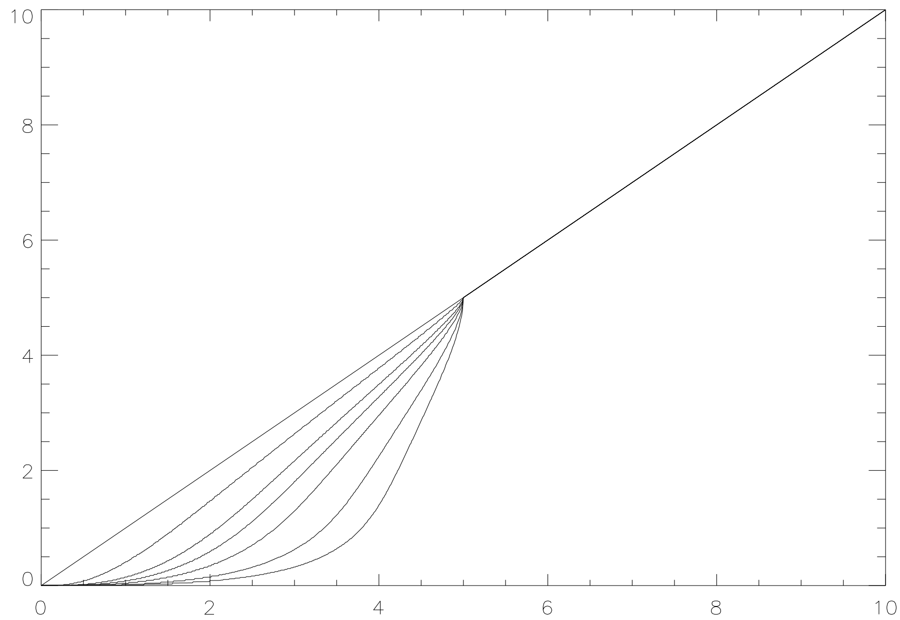

Figure 3.2: Coefficient d'ondelettes corrigé en fonction du coefficient d'ondelettes pour différentes valeurs de α.

paramètre finalement utilisé est choisi comme étant le produit d'un paramètre utilisateur (pris par défaut à 1) et la valeur calculée α_c. Ceci permet à l'utilisateur de garder un contrôle sur la possibilité d'introduire un sur-lissage, ou un sous-lissage. Un tel algorithme est nécessairement itératif, et ça prend donc plus de temps qu'un simple seuillage dur.

- On peut encore introduire plus de contraintes sur α en utilisant le fait que le résidu devrait avoir un écart type de bruit à chaque échelle j égal à l'écart type du bruit σ_j à la même échelle. Alors, on aura un α_j par échelle.

Une autre manière de fixer α est d'introduire une distribution a priori (ou une connaissance a priori) sur la manière dont le filtrage devrait fonctionner. Par exemple, dans le cas de la restauration d'images astronomiques, l'utilisateur préfère en général que le flux (intensité totale) contenu dans une étoile ou une galaxie ne soit pas modifié au cours du processus de restauration. Ceci signifie que le résidu aux positions des objets astronomiques sera à peu près nul. Toutes les zones à zéro dans la carte des résidus ne représentent clairement pas un comportement réaliste du bruit, mais du point de vue de l'utilisateur, il est essentiel que les objets visibles dans l'image filtrée contiennent le même flux que dans les données brutes. Pour avoir ce type de régularisation, le paramètre α n'est plus une constante, mais dépend des données brutes. Ainsi, on a un α par coefficient d'ondelettes que l'on écrit $\alpha_s(w_{j,k})$, et qui peut obtenu par

$$\alpha_s(w_{j,k}) = \alpha_j \frac{1 - L(w_{j,k})}{L(w_{j,k})} \tag{3.10}$$

où $L(w_{j,k}) = MIN(1, \frac{|w_{j,k}|}{k_s \sigma_j})$, et k_s est un paramètre utilisateur (en général pris égal à 3).

Quand $L(w_{j,k})$ est proche de 1, $\alpha_s(w_{j,k})$ devient nul, et il n'y a plus de régularisation, et une solution évidente $\tilde{w}_{j,k} = w_{j,k}$. De cette manière, le coefficient d'ondelettes est préservé de la régularisation. Si $L(w_{j,k})$ est proche de 0, $\alpha_s(w_{j,k})$ tend vers l'infini, et le premier terme dans l'équation (3.9) est négligeable, et la solution est $\tilde{w}_{j,k} = 0$. En pratique, tous les coefficients plus grands que $k_s\sigma_j$ ne sont pas modifiés comme dans le seuillage dur. On note aussi qu'en prenant une distribution de $L(w_{j,k})$ égale à 0 ou 1 (1 si $\mid w_{j,k} \mid > k\sigma$ par exemple), la solution est la même que le seuillage dur. Introduire le SNR dans le paramètre α n'est pas incompatible avec le fait que l'on cherche un résidu qui vérifie les propriétés d'un bruit, et on peut toujours chercher un α_j par échelle.

L'utilisation d'un modèle

Si on a un modèle D_m pour les données, il peut être naturellement introduit dans les équations de filtrage (Starck and Murtagh, 1999):

$$J_m(\tilde{D}) = H_s(D - \tilde{D}) + \alpha H_n(\tilde{D} - D_m) \tag{3.11}$$

ou, pour chaque coefficient w:

$$j_m(\tilde{w}_{j,k}) = h_s(w_{j,k} - \tilde{w}_{j,k}) + \alpha h_n(\tilde{w}_{j,k} - w^m_{j,k}) \tag{3.12}$$

$w^m_{j,k}$ est le coefficient d'ondelettes correspondant de D_m.

Le modèle peut être de différents types. Il peut être une image, et dans ce cas, les coefficients $w^m_{j,k}$ sont obtenus par une simple transformée en ondelettes de l'image. Il peut aussi être exprimé par une distribution d'une fonction donnée qui fournit le coefficient d'ondelettes modèle w_m à partir des données. Par exemple, le cas où l'on veut garder intact les coefficients d'ondelettes de fort rapport signal sur bruit (voir l'équation 3.10) peut aussi être traité par l'utilisation d'un modèle, simplement en calculant $w^m_{j,k}$ par

$$w^m_{j,k} = P_s(w_{j,k}) w_{j,k} \tag{3.13}$$

Si $w_{j,k}$ a un fort SNR, $P_s(w_{j,k})$ et proche de 1, et $w^m_{j,k}$ est égal à $w_{j,k}$. Alors $\alpha h_n(\tilde{w}_{j,k} - w^m_{j,k})$ est égal à zéro et $\tilde{w}_{j,k} = w_{j,k}$, c'est à dire qu'aucune régularisation n'est effectuée sur $w_{j,k}$.

D'autres modèles peuvent bien évidemment être utilisés. Si une image contient des contours, il peut être intéressant de dériver le modèle à partir des contours détectés.

Si on une information a priori sur le contenu de l'image, elle doit être utilisée pour améliorer la qualité du filtrage.

3.1.4 L'algorithme de filtrage par entropie multiéchelle

L'algorithme de filtrage par entropie multiéchelle (MEF) (Starck and Murtagh, 1999) consiste à minimiser chaque coefficient d'ondelettes $w_{j,k}$ à l'échelle j la quantité

$$j_m(\tilde{w}_{j,k}) = h_s(w_{j,k} - \tilde{w}_{j,k}) + \alpha_j h_n(\tilde{w}_{j,k} - w^m_{j,k}) \tag{3.14}$$

ou

$$j_{ms}(\tilde{w}_{j,k}) = h_s(w_{j,k} - \tilde{w}_{j,k}) + \alpha_j \alpha_s(w_{j,k}) h_n(\tilde{w}_{j,k} - w^m_{j,k}) \tag{3.15}$$

si le SNR est utilisé. Par défaut, le modèle $w^m_{j,k}$ est égal à 0. Il n'y a pas de paramètre utilisateur car α_j est calculé automatiquement pour que le résidu vérifie bien les propriétés du bruit. Si un sur-lissage (ou un sous-lissage) est désiré, un paramètre utilisateur doit être introduit. On propose dans ce cas de calculer α_j comme auparavant, puis de le multiplier par une valeur α_u donné par l'utilisateur (la valeur de α_u est donc égale à 1). En augmentant α_u, on a alors un sur-lissage, tandis qu'en l'abaissant, on a au contraire un sous-lissage.

En utilisant une simple dichotomie, l'algorithme est:

1. Estimer le niveau de bruit dans les données σ (voir (Olsen, 1993; Starck and Murtagh, 1998)).
2. Calculer la transformée en ondelettes des données.
3. Calculer à partir de σ le niveau de bruit σ_j à chaque échelle j.
4. Initialiser $\alpha_j^{min} = 0$, $\alpha_j^{max} = A_{\max}$. $A_{\max}$ doit être suffisemment grand pour encadrer la solution. Une valeur de 200 a été utilisée dans notre programme.
5. Pour chaque échelle j faire
   (a) Initialiser $\alpha_j = \frac{\alpha_j^{min}+\alpha_j^{max}}{2}$
   (b) Pour chaque coefficient $w_{j,k}$ de l'échelle j, trouver $\tilde{w}_{j,k}$ en minimisant $j_m(\tilde{w}_{j,k})$ ou $j_{ms}(\tilde{w}_{j,k})$
   (c) Calculer l'écart type du résidu: $\sigma_j^r = \sqrt{\frac{1}{N_j}\sum_{k=1}^{N_j}(w_{j,k} - \tilde{w}_{j,k})^2}$
   (d) si $\sigma_j^r > \sigma_j$ alors la régularisation est trop forte, et on fixe α_j^{max} à α_j, sinon on fixe α_j^{min} à α_j.
6. Si $\alpha_j^{max} - \alpha_j^{min} > \epsilon$ alors aller à 5.
7. Multiplier tous les α_j par la constante α_u permettant à l'utilisateur d'affiner le lissage s'il le souhaite.
8. Pour chaque échelle j pour chaque coefficient $w_{j,k}$ trouver $\tilde{w}_{j,k}$ en minimisant $j_m(\tilde{w}_{j,k})$ ou $j_{ms}(\tilde{w}_{j,k})$.
9. Reconstruire les données filtrées à partir $\tilde{w}_{j,k}$ par une transformée ondelette inverse.

La minimisation de j_m, ou de j_{ms}, (étape 5b) peut être faite par n'importe quelle méthode. Par exemple, une simple dichotomie peut être utilisée pour trouver $\tilde{w}$ tel que

$$\nabla(j_m(\tilde{w}_{j,k})) = \frac{\partial h_s(w_{j,k} - \tilde{w}_{j,k})}{\partial \tilde{w}_{j,k}} + \alpha_j \frac{\partial h_n(\tilde{w}_{j,k})}{\partial \tilde{w}_{j,k}} = 0 \tag{3.16}$$

Dans le cas d'un bruit Gaussien et en utilisant l'approche N2-MSE, on trouve, à partir des équations 3.49 et 3.54 décrites dans la section suivante, que $\nabla(j_m(\tilde{w}_{j,k}))$ est défini par:

$$\begin{aligned}\nabla(j(\tilde{w}_{j,k})) = & -\frac{w_{j,k} - \tilde{w}_{j,k}}{\sigma_j^2}\mathrm{erf}\left(\frac{w_{j,k} - \tilde{w}_{j,k}}{\sqrt{2}\sigma_j}\right) + \sqrt{\frac{2}{\pi}}\frac{1}{\sigma_j}\left[1 - e^{-\frac{(w_{j,k}-\tilde{w}_{j,k})^2}{2\sigma_j^2}}\right] + \\ & \alpha_j\left(\frac{\tilde{w}_{j,k}}{\sigma_j^2}\mathrm{erfc}\left(\frac{\tilde{w}_{j,k}}{\sqrt{2}\sigma_j}\right) + \frac{1}{\sigma_j}\sqrt{\frac{2}{\pi}}\left[1 - e^{-\frac{\tilde{w}_{j,k}^2}{2\sigma_j^2}}\right]\right)\end{aligned} \tag{3.17}$$

3.1.5 Calcul des dérivées de h_s et h_n dans le cas d'un bruit Gaussien

La fonction erf

La fonction erf(x), complémentaire de la fonction erfc(x) est défini par

$$\mathrm{erf}(\mathrm{x}) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \tag{3.18}$$

$$\mathrm{erfc}(x) = 1 - \mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt \tag{3.19}$$

Ces fonctions ont les limites et symétries suivantes

erf(0) = 0	erf(∞) = 1
erfc(0) = 1	erfc(∞) = 0
erf(-x) = erf(x)	erfc(-x) = 2-erfc(x)

(3.20)

Approche N1-MSE

La contribution d'un coefficient d'ondelette x à l'information bruit est

$$h_n(x) = \frac{x^2}{2\sigma^2} \mathrm{erfc}(\frac{x}{\sqrt{2}\sigma}) \tag{3.21}$$

$$\frac{dh_n(x)}{dx} = \frac{x}{\sigma^2} \mathrm{erfc}(\frac{x}{\sqrt{2}\sigma}) + \frac{x^2}{2\sigma^2} \frac{\partial \mathrm{erfc}(\frac{x}{\sqrt{2}\sigma})}{\partial x} \tag{3.22}$$

On dérive la fonction erfc:

$$\frac{\partial\, \mathrm{erfc}(x)}{\partial x} = -\frac{2}{\sqrt{\pi}} e^{-x^2} \tag{3.23}$$

alors

$$\frac{dh_n(x)}{dx} = \frac{x}{\sigma^2} \mathrm{erfc}(\frac{x}{\sqrt{2}\sigma}) - \frac{x^2}{\sqrt{2\pi}\sigma^3} e^{-\frac{x^2}{2\sigma^2}} \tag{3.24}$$

Dans le but de minimiser la fonctionnelle 3.12, on veut calculer calculer les dérivées de $h_s(y-x)$, $h_s(y-x)$ mesurant le montant d'information contenu dans le résidu (y étant les données).

$$h_s(x-y) = \frac{(y-x)^2}{2\sigma^2} \mathrm{erf}(\frac{y-x}{\sqrt{2}\sigma}) \tag{3.25}$$

en notant $z = y - x$, on a

$$h_s(z) = \frac{z^2}{2\sigma^2} \mathrm{erf}(\frac{z}{\sqrt{2}\sigma}) \tag{3.26}$$

$$\begin{aligned}
\frac{dh_s(z)}{dz} &= \frac{z}{\sigma^2}\mathrm{erf}(\frac{z}{\sqrt{2}\sigma}) + \frac{z^2}{2\sigma^2}\frac{\partial \mathrm{erf}(\frac{z}{\sqrt{2}\sigma})}{\partial z} \qquad (3.27)\\
&= \frac{z}{\sigma^2}\mathrm{erf}(\frac{z}{\sqrt{2}\sigma}) + \frac{z^2}{\sqrt{2\pi}\sigma^3}e^{-\frac{z^2}{2\sigma^2}} \qquad (3.28)
\end{aligned}$$

et

$$\frac{dh_s(y-x)}{dx} = -\frac{dh_s(z)}{dz} \tag{3.29}$$

Approche N2-MSE

La contribution d'un coefficient d'ondelette x à l'information bruit est

$$h_n(x) = \frac{1}{\sigma^2}\int_0^x t\,\mathrm{erfc}(\frac{x-t}{\sqrt{2}\sigma})dt \tag{3.30}$$

$$\begin{aligned}
\frac{dh_n(x)}{dx} &= h_n(x+dx) - h_n(x)\\
&= \frac{1}{\sigma^2}\int_0^{x+dx}\mathrm{erfc}(\frac{x+dx-t}{\sqrt{2}\sigma})dt - \frac{1}{\sigma^2}\int_0^x \mathrm{erfc}(\frac{x-t}{\sqrt{2}\sigma})dt \qquad (3.31)
\end{aligned}$$

$$\frac{dh_n(x)}{dx} = \frac{x}{\sigma^2}\mathrm{erfc}(\frac{x-x}{\sqrt{2}\sigma}) + \frac{1}{\sigma^2}\int_0^x \frac{\partial\, t\,\mathrm{erfc}(\frac{x-t}{\sqrt{2}\sigma})}{\partial x}dt \tag{3.32}$$

comme $\mathrm{erfc}(0) = 1$ on a:

$$\frac{dh_n(x)}{dx} = \frac{x}{\sigma^2} + \frac{1}{\sigma^2}\int_0^x \frac{\partial\, t\,\mathrm{erfc}(\frac{x-t}{\sqrt{2}\sigma})}{\partial x}dt \tag{3.33}$$

On dérive la fonction erfc:

$$\frac{\partial\,\mathrm{erfc}(x)}{\partial x} = -\frac{2}{\sqrt{\pi}}e^{-x^2} \tag{3.34}$$

$$\frac{\partial\,\mathrm{erfc}(\frac{(x-t)}{\sqrt{2}\sigma})}{\partial x} = -\frac{2}{\sqrt{\pi}}\frac{1}{\sqrt{2}\sigma}e^{-\frac{(x-t)^2}{2\sigma^2}} = -\sqrt{\frac{2}{\pi}}\frac{1}{\sigma}e^{-\frac{(x-t)^2}{2\sigma^2}} \tag{3.35}$$

On en déduit pour la dérivée de h_n:

$$\frac{dh_n(x)}{dx} = \frac{x}{\sigma^2} + \frac{1}{\sigma^2}\int_0^x -\sqrt{\frac{2}{\pi}}\frac{1}{\sigma}t\,e^{-\frac{(x-t)^2}{2\sigma^2}}dt \tag{3.36}$$

$$\frac{dh_n(x)}{dx} = \frac{x}{\sigma^2} + \frac{1}{\sigma^3}\sqrt{\frac{2}{\pi}}\int_0^x t\,e^{-\frac{(x-t)^2}{2\sigma^2}}dt \tag{3.37}$$

On défini la variable J par

$$J = \int_0^x t\, e^{-\frac{(x-t)^2}{2\sigma^2}} dt \tag{3.38}$$

ainsi que u par

$$\begin{array}{rr} u = \frac{t-x}{\sqrt{2}\sigma} & t = x + u\sqrt{2}\sigma \\ & dt = \sqrt{2}\sigma du \\ t = 0 \Rightarrow u = \frac{-x}{\sqrt{2}\sigma} & t = x \Rightarrow u = 0 \end{array} \tag{3.39}$$

J peut être écrit en fonction de u

$$J = \int_{\frac{-x}{\sqrt{2}\sigma}}^{0} (x + u\sqrt{2}\sigma) e^{-u^2} \sqrt{2}\sigma du \tag{3.40}$$

$$J = \sqrt{2}\sigma x \int_{\frac{-x}{\sqrt{2}\sigma}}^{0} e^{-u^2} du + 2\sigma^2 \int_{\frac{-x}{\sqrt{2}\sigma}}^{0} u\, e^{-u^2} du \tag{3.41}$$

Le premier terme de J peut s'écrire sous la forme:

$$J_0 = \sqrt{2}\sigma x \int_0^{\frac{x}{\sqrt{2}\sigma}} e^{-u^2} du \tag{3.42}$$

En exprimant J_0 avec la fonction erf.

$$J_0 = \sqrt{2}\sigma \frac{\sqrt{\pi}}{2} x \operatorname{erf}(\frac{x}{\sqrt{2}\sigma}) J_0 = \sigma\sqrt{\frac{\pi}{2}}\, x \operatorname{erf}(\frac{x}{\sqrt{2}\sigma}) \tag{3.43}$$

Le second terme de J est

$$J_1 = 2\sigma^2 \int_{\frac{-x}{\sqrt{2}\sigma}}^{0} u\, e^{-u^2} du \tag{3.44}$$

avec

$$\frac{de^{-u^2}}{du} = -2ue^{-u^2} \tag{3.45}$$

donc

$$J_1 = -\sigma^2 \int_{\frac{-x}{\sqrt{2}\sigma}}^{0} d(e^{-u^2}) \tag{3.46}$$

$$J_1 = \sigma^2 [e^{-\frac{x^2}{2\sigma^2}} - 1] \tag{3.47}$$

On en déduit que

$$J = J_0 + J_1 = \sigma\sqrt{\frac{\pi}{2}}\, x \operatorname{erf}(\frac{x}{\sqrt{2}\sigma}) + \sigma^2 [e^{-\frac{x^2}{2\sigma^2}} - 1] \tag{3.48}$$

et la dérivée de h_n est

$$\begin{aligned}
\frac{dh_n(x)}{dx} &= \frac{x}{\sigma^2} - \frac{1}{\sigma^3}\sqrt{\frac{2}{\pi}}\,J \\
&= \frac{x}{\sigma^2} - \frac{x}{\sigma^2}\,x\,\mathrm{erf}(\frac{x}{\sqrt{2}\sigma}) + \frac{1}{\sigma}\sqrt{\frac{2}{\pi}}\,[1 - e^{-\frac{x2}{2\sigma 2}}] \\
&= \frac{x}{\sigma^2}\mathrm{erfc}(\frac{x}{\sqrt{2}\sigma}) + \frac{1}{\sigma}\sqrt{\frac{2}{\pi}}\,[1 - e^{-\frac{x^2}{2\sigma^2}}] \qquad (3.49)
\end{aligned}$$

Dans le but de minimiser la fonctionnelle 3.7, il faut aussi calculer la dérivée de $h_s(y-x)$, $h_s(y-x)$ étant l'information signal du résidu

$$h_s(y-x) = \frac{1}{\sigma^2}\int_0^{y-x} t\ \mathrm{erf}(\frac{y-x-t}{\sqrt{2}\sigma})dt \qquad (3.50)$$

En effectuant le changement de variable $z = y - x$, on a

$$\begin{aligned}
h_s(z) &= \frac{1}{\sigma^2}\int_0^{z} t\ \mathrm{erf}(\frac{z-t}{\sqrt{2}\sigma})dt \\
&= \frac{1}{\sigma^2}\int_0^{z} t dt - \frac{1}{\sigma^2}\int_0^{z} t\ \mathrm{erfc}(\frac{z-t}{\sqrt{2}\sigma})dt \qquad (3.51)
\end{aligned}$$

et

$$\frac{dh_s(y-x)}{dx} = \frac{\partial h_s(z)}{\partial z}\frac{\partial z}{\partial x} \qquad (3.52)$$

$$\frac{\partial h_s(z)}{\partial z} = \frac{z}{\sigma^2} - \frac{\partial h_n(z)}{\partial z} \qquad (3.53)$$

alors

$$\begin{aligned}
\frac{dh_s(y-x)}{dx} &= -\frac{y-x}{\sigma^2} + \frac{y-x}{\sigma^2}\mathrm{erfc}(\frac{y-x}{\sqrt{2}\sigma}) + \sqrt{\frac{2}{\pi}}\frac{1}{\sigma}[1 - e^{-\frac{(y-x)^2}{2\sigma^2}}] \\
&= -\frac{y-x}{\sigma^2}\mathrm{erf}(\frac{y-x}{\sqrt{2}\sigma}) + \sqrt{\frac{2}{\pi}}\frac{1}{\sigma}[1 - e^{-\frac{(y-x)^2}{2\sigma^2}}] \qquad (3.54)
\end{aligned}$$

3.1.6 Optimisation

Dans le cas d'un bruit Gaussien, le calcul des fonctions erf et $erfc$ peut conduire à un important temps de calcul. Ceci peut être évité en pré-calculcant des tables, ce qui est possible grâce aux propriétés de $\frac{\partial h_s}{\partial \tilde{w}}$ et $\frac{\partial h_n}{\partial \tilde{w}}$. h_s et h_n sont des fonctions de l'écart type du bruit, et on définit les fonctions réduites h_s^r et h_n^r comme étant h_s et h_n quand l'écart type du bruit est égal à 1. Il est facile de vérifier que

$$\frac{\partial h_s(w_{j,k})}{\partial \tilde{w}} = \sigma_j \frac{\partial h_s^r(\frac{w_{j,k}}{\sigma_j})}{\partial \tilde{w}} \qquad (3.55)$$

$$\frac{\partial h_n(w_{j,k})}{\partial \tilde{w}} = \sigma_j \frac{\partial h_n^r(\frac{w_{j,k}}{\sigma_j})}{\partial \tilde{w}} \qquad (3.56)$$

De plus, $\frac{\partial h_n^r}{\partial \tilde{w}}$ et $\frac{\partial h_s^r}{\partial \tilde{w}}$ sont des fonctions symétriques, $\frac{\partial h_n^r}{\partial \tilde{w}}$ et convergent vers une valeur constante C (C=0.798), et $\frac{\partial h_s^r}{\partial \tilde{w}}$ tend vers $C - w$ quand w est suffisamment grand (> 5). Dans notre code, on pré-calcule les tables en utilisant un pas de 0.01 entre 0 et 5. Si aucun modèle n'est introduit, et si le SNR n'est pas utilisé, le coefficient d'ondelettes filtré est une fonction de α et de $\frac{w_{j,k}}{\sigma_j}$, et un second niveau d'optimisation peut être introduit en pré-calculant des tables de solutions pour différentes valeurs de α.

3.1.7 Exemples

Filtrage de données 1D

Les figures 3.3, 3.4, et 3.5 montrent les résultats après filtrage par l'entropie multiéchelle. Sur les figures 3.3 et 3.4, on peut voir en (1a) les données simulées, en (1b) les données bruitées, en (2a) les données filtrées, en (2b) le résidu, en (3a) les données bruités en pointillé avec les données filtrées superposées en continu sur la partie [500,1000], (3b) à gauche le résidu sur la même partie, en (4) même chose qu'en (3) mais sur la partie [1500,2000]. La figure 3.5 montre respectivement de haut en bas les données simulées, les données bruitées, les données filtrées, et le résidu.

Pour les filtrages de la figure 3.3 et 3.5, on a utilisé l'option de protection des coefficients basée sur le SNR (protection à 5σ), les autres paramètres étant pris par défault (l'écart type du bruit et α_j sont automatiquement calculés, $\alpha_u = 1$, l'algorithme de transformée en ondelettes est l'algorithme à trous). Pour le signal composé de blocks (Fig. 3.4), on a utilisé la transformée médiane multirésolution à la place de l'algorithme à trous.

La figure 3.6 montre le filtrage d'un vrai spectre. La dernière courbe représente la différence entre les données et les données filtrées. On peut voir que le résidu ne contient que du bruit. Là encore, on n'a pas utilisé de paramètre particulier, mais on a utilisé l'option de protection des coefficients basée sur le SNR, afin de bien restaurer le flux des diverses bandes.

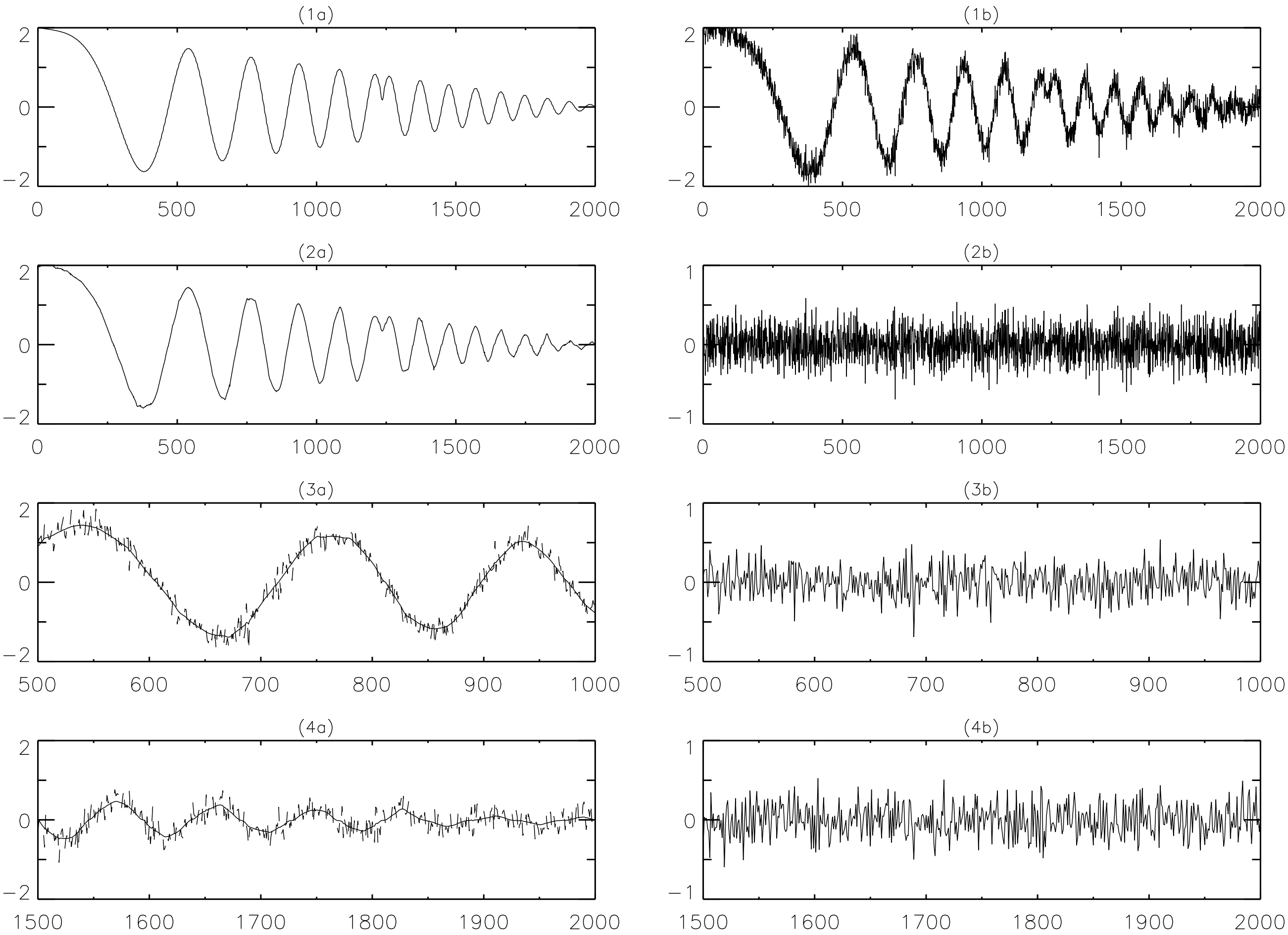

Figure 3.3: De haut en bas, (1a) les données simulées et (1b) les données bruitées, (2a) les données filtrées et (2b) le résidu, (3a) les données bruités en pointillé avec les données filtrées superposées en continu sur la partie [500,1000], et (3b) à gauche le résidu sur la même partie, (4) même chose que (3), mais sur la partie [1500,2000].

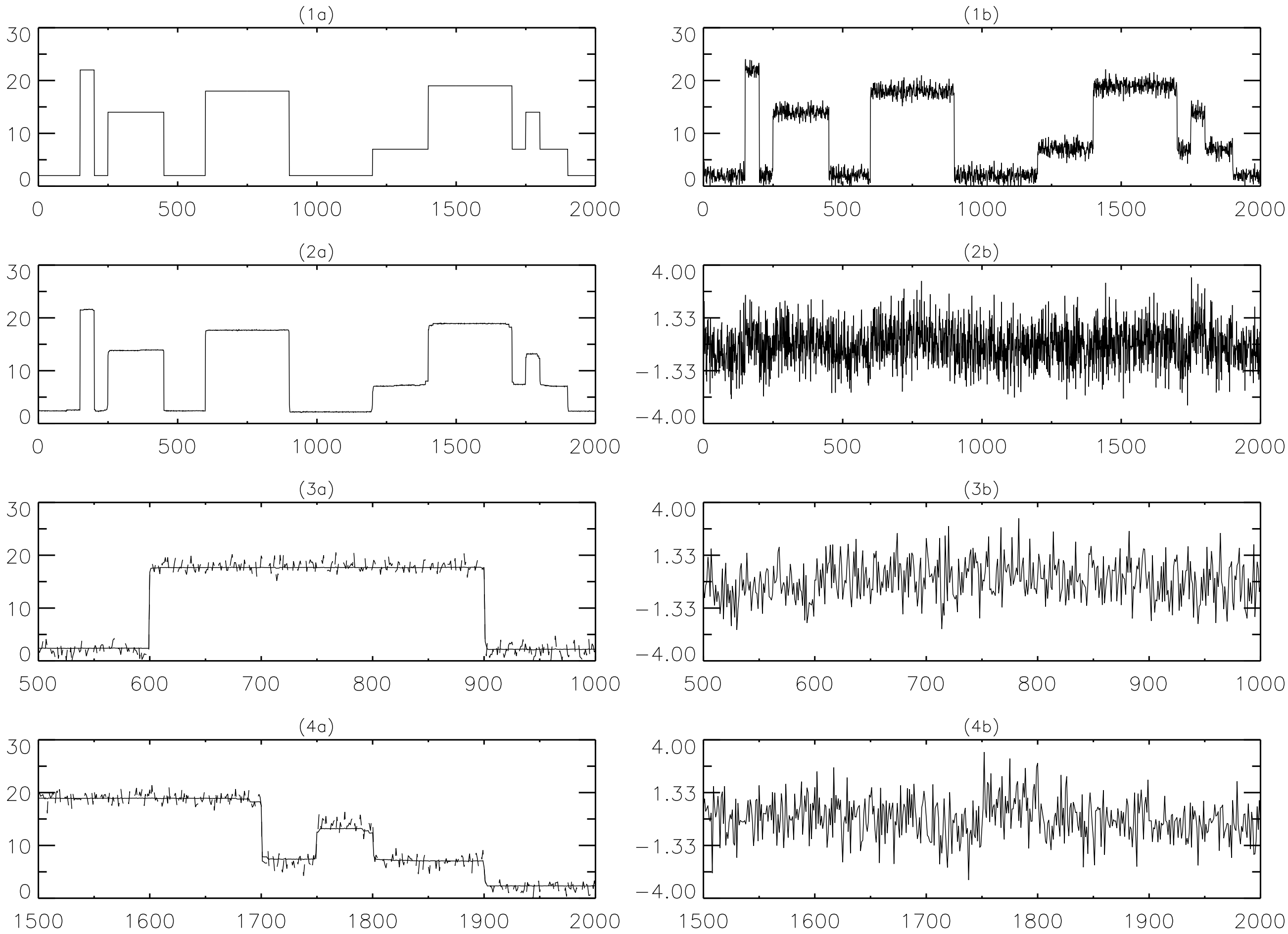

Figure 3.4: De haut en bas, (1a) les données simulées et (1b) les données bruitées, (2a) les données filtrées et (2b) le résidu, (3a) les données bruités en pointillé avec les données filtrées superposées en continu sur la partie [500,1000], et (3b) à gauche le résidu sur la même partie, (4) même chose que (3), mais sur la partie [1500,2000].

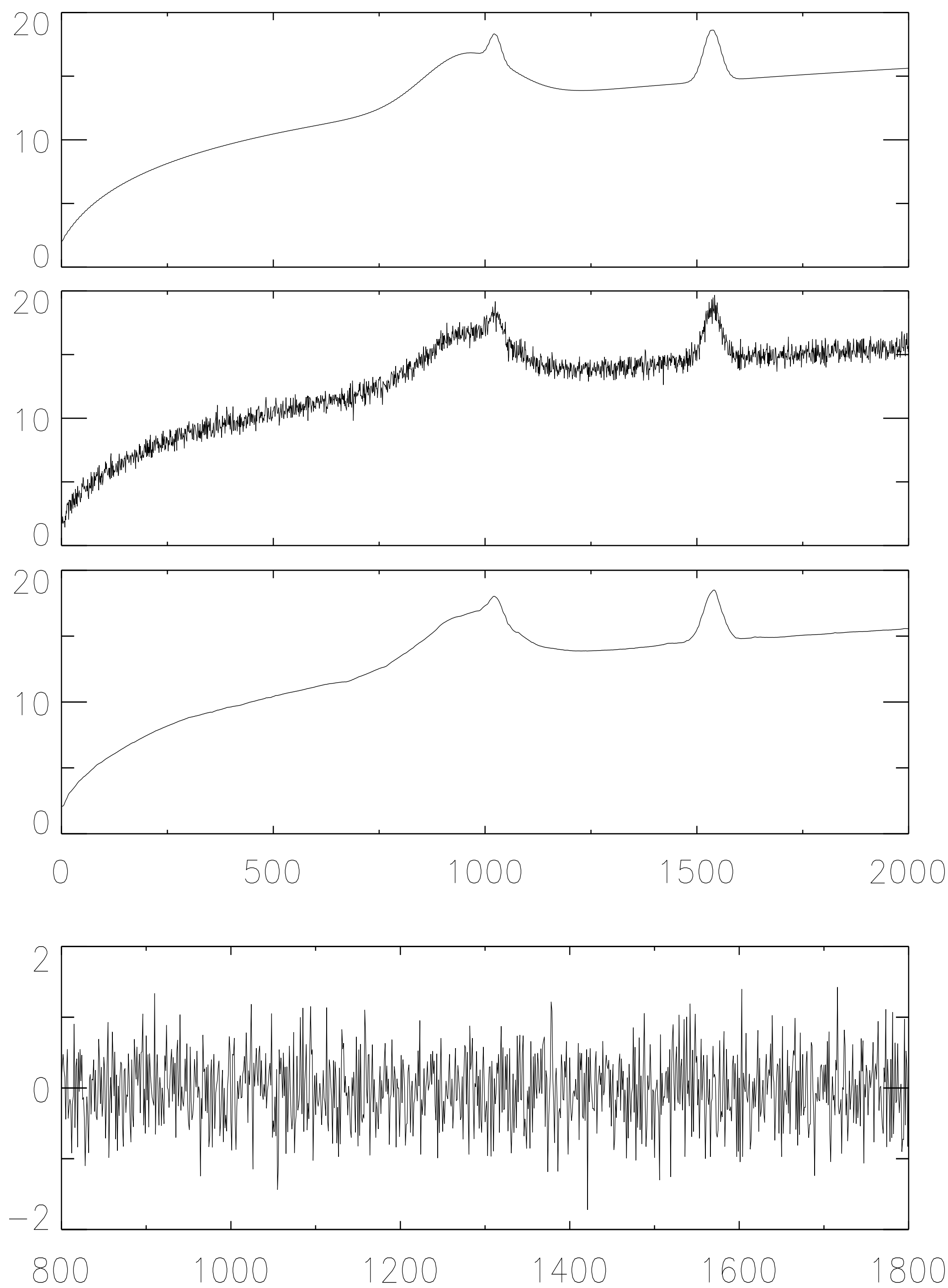

Figure 3.5: De haut en bas, les données simulées, les données bruitées, les données filtrées, et le résidu sur la partie [800,1800].

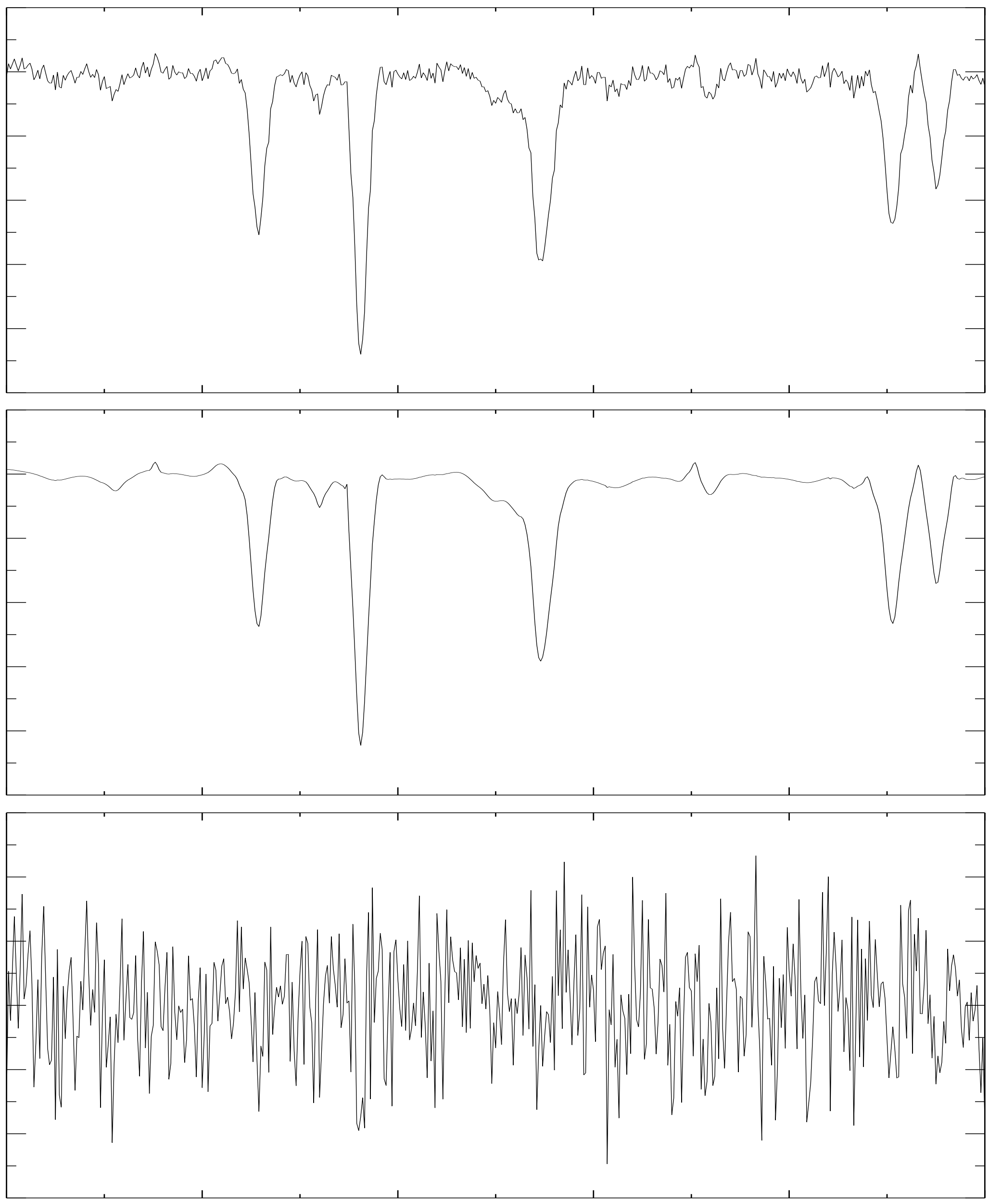

Figure 3.6: De haut en bas, spectre réel, spectre filtré, et différence entre le spectre et les données filtrées. Comme on peut le voir, aucune structure significative n'apparaît dans le résidu.

Filtrage d'images

Une simulation est présentée figure 3.7. On peut voir en a,b,c,d respectivement une image simulée contenant des étoiles et des galaxies, la même image mais bruitée (bruit Gaussien), l'image filtrée, et la différence entre l'image bruitée et l'image filtrée. On peut voir que le résidu ne contient aucune structure.

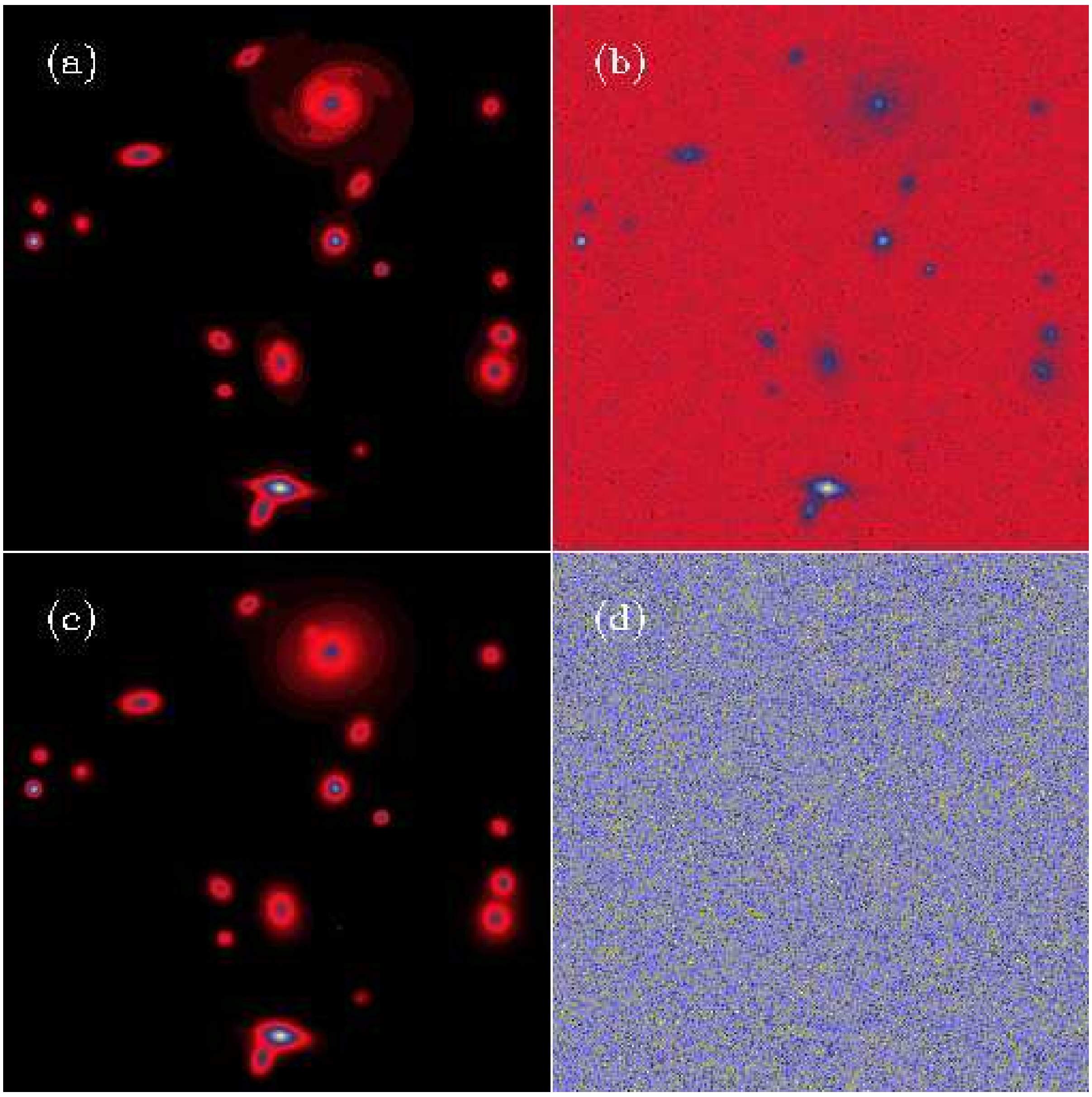

Figure 3.7: (a) Image simulée, (b) image simulée + bruit Gaussien, (c) image filtrée, et (d) résidu (image (b) - image (c)).

3.1.8 Comparaison à partir de simulations

Descriptions des simulations

Une série de simulations a été réalisées à partir de deux images: l'image 512x512 classique de Lena, et une image 512x512 d'un paysage. Pour chacune de ces images, on a créé trois autres images en générant des bruits gaussiens d'écart type 5,10,30. Les six images simulées ont ensuite été filtrées en utilisant différentes transformées multiéchelles, et différentes méthodes de correction du bruit. Les transformées testées sont

1. la transformée en ondelettes de Haar (FWT-Haar).
2. la transformée en ondelettes bi-orthogonales de Mallat-Daubechies avec les filtres 7/9 (Antonini et al., 1992) (FWT-7/9).
3. la transformée en ondelettes de Feauveau.
4. l'algorithme à trous en prenant une B-spline pour la fonction d'échelle (voir (Starck et al., 1998a) pour plus de détails).
5. la transformée médiane multirésolution (MMT) (Starck et al., 1996e).

Les deux premières appartiennent à la classe des transformées rapides. La troisième est une transformée non redondante, mais utilise une fonction ondelette isotrope. L'algorithme à trous est redondant avec une fonction ondelette isotrope. La transformée MMT n'est pas une transformée en ondelette, mais permet une représentation multirésolution.

En utilisant ces cinq transformées, on a testé huit stratégies différentes pour corriger les coefficients d'ondelettes du bruit:

- seuillage k-sigma dur et doux
- seuillage de Donoho dur et doux
- méthode entropie multiéchelle
- seuillage hiérarchique dur
- filtre de Wiener multirésolution
- filtre de Wiener hiérarchique

Les trois dernières stratégies n'ont été développées que pour des transformées redondantes (algorithme à trous et MMT dans notre cas).

Finalement, environ 200 images filtrées ont été crées. Quatre niveaux de résolution ont été utilisés pour le filtrage, la constante k pour le seuillage dur a toujours été choisi à 4 pour la première échelle, et 3 pour les suivantes. Pour la méthode entropie multiéchelle, le paramètre α est déterminé automatiquement par le code afin d'obtenir un résidu avec un écart type égal à l'écart type du bruit.

Pour chaque image filtrée, l'erreur normalisée entre l'image originale I et l'image filtrée $\tilde{I}$ a été calculée par:

$$\mathbf{E}^2_{\mathrm{nms}} = \frac{\sum_{pix}(I - \tilde{I})^2}{\sum_{pix} I^2} \tag{3.57}$$

et le rapport **PSNR** (*peak signal-to-noise*) est:

$$\mathbf{PSNR}_{\mathrm{dB}} = 10 \log_{10} \frac{255}{\mathbf{E}^2_{\mathrm{nms}}} \tag{3.58}$$

On a aussi calculé le facteur de corrélation, mais on a trouvé qu'il n'apportait pas plus d'information que le PSNR. Si le PSNR est une valeur objective, il ne rend pas compte de l'aspect visuel de l'image. On n'a donc aussi visualisé toutes ces images pour indiquer si des artéfacts sont visibles.

Les résultats de ces simulations sont présentés dans les tables 3.1 à 3.6.

Méthode	FWT-Haar	FWT-7/9	Feauveau	à trous	MMT
Seuil. dur.	34.63	35.95	33.27	35.20	34.82
Seuil. doux.	32.35	33.83	30.67	32.30	32.43
Seuil. dur Donoho	33.19	34.62	31.05	33.98	33.68
Seuil. doux Donoho	30.69	32.09	28.73	30.76	31.19
Seuil. hiérarchique	-	-	-	35.26	34.89
Wiener hiérarchique	-	-	-	33.35	31.91
Wiener multirésol.	-	-	-	33.42	31.93
Entropie multiéchelle	35.86	36.76	-	35.82	35.56

Table 3.1: PSNR après filtrage de l'image simulée (Lena + bruit Gaussien (sigma=5)).

Méthode	FWT-Haar	FWT-7/9	Feauveau	à trous	MMT
Seuil. dur.	31.31	32.97	29.87	32.63	31.80
Seuil. doux.	29.72	31.29	28.05	30.03	30.15
Seuil. dur Donoho	29.94	31.55	27.68	31.33	30.88
Seuil. doux Donoho	28.18	29.66	26.77	28.49	29.09
Seuil. hiérarchique	-	-	-	32.75	31.93
Wiener hiérarchique	-	-	-	31.71	30.33
Wiener multirésol.	-	-	-	31.68	30.24
Entropie multiéchelle	32.12	33.39	-	32.41	31.95

Table 3.2: PSNR après filtrage de l'image simulée (Lena + bruit Gaussien (sigma=10)).

Méthode	FWT-Haar	FWT-7/9	Feauveau	à trous	MMT
Seuil. dur.	26.82	27.97	26.00	28.58	28.19
Seuil. doux.	26.27	27.67	25.85	26.85	27.27
Seuil. dur Donoho	25.99	27.46	25.80	27.03	27.42
Seuil. doux Donoho	25.29	26.78	25.80	25.85	26.58
Seuil. hiérarchique	-	-	-	28.97	28.42
Wiener hiérarchique	-	-	-	28.08	27.96
Wiener multirésol.	-	-	-	27.25	26.81
Entropie multiéchelle	27.45	28.75	-	28.37	27.96

Table 3.3: PSNR après filtrage de l'image simulée (Lena + bruit Gaussien (sigma=30)).

Méthode	FWT-Haar	FWT-7/9	Feauveau	à trous	MMT
Seuil. dur.	32.50	33.02	30.48	32.49	31.79
Seuil. doux.	30.35	30.97	28.27	29.87	29.59
Seuil. dur Donoho	31.04	31.53	28.38	31.23	30.67
Seuil. doux Donoho	28.80	29.40	26.70	28.46	28.45
Seuil. hiérarchique	-	-	-	32.51	31.82
Wiener hiérarchique	-	-	-	30.59	30.32
Wiener multirésol.	-	-	-	30.65	30.35
Entropie multiéchelle	34.63	34.94	-	34.30	33.97

Table 3.4: PSNR après filtrage de l'image simulée (paysage + bruit Gaussien (sigma=5)).

Méthode	FWT-Haar	FWT-7/9	Feauveau	à trous	MMT
Seuil. dur.	29.32	30.00	27.38	29.88	28.91
Seuil. doux.	27.89	28.66	26.18	27.78	27.45
Seuil. dur Donoho	28.05	28.74	25.86	28.58	27.98
Seuil. doux Donoho	26.53	27.32	25.32	26.50	26.52
Seuil. hiérarchique	-	-	-	29.99	28.99
Wiener hiérarchique	-	-	-	29.59	28.04
Wiener multirésol.	-	-	-	29.64	28.04
Entropie multiéchelle	30.80	31.35	-	30.70	30.16

Table 3.5: PSNR après filtrage de l'image simulée (paysage + bruit Gaussien (sigma=10)).

Méthode	FWT-Haar	FWT-7/9	Feauveau	à trous	MMT
Seuil. dur.	25.44	26.01	24.80	26.55	25.90
Seuil. doux.	25.03	25.88	24.75	25.33	25.10
Seuil. dur Donoho	24.77	25.60	24.72	25.36	25.19
Seuil. doux Donoho	24.26	25.29	24.725	24.61	24.51
Seuil. hiérarchique	-	-	-	27.07	26.21
Wiener hiérarchique	-	-	-	26.52	25.83
Wiener multirésol.	-	-	-	25.89	25.24
Entropie multiéchelle	26.33	27.11	-	26.88	26.16

Table 3.6: PSNR après filtrage de l'image simulée (paysage + bruit Gaussien (sigma=30)).

L'analyse des simulations

L'algorithme de transformée multiéchelle

Le filtrage par transformée de Haar produit toujours des artéfacts, quelque soit le niveau de bruit. En utilisant d'autres types de filtres, les artéfacts n'apparaissent seulement qu'au delà d'un certain niveau de bruit. Améliorer le jeu de filtres améliore la qualité de l'image filtrée, ce qui est un résultat bien connu. Quand le bruit augmente, des artéfacts apparaissent, même avec de bons filtres comme ceux d'Antonini (filtres 7/9) (1992).

- **Transformée en ondelettes de Feauveau.** La transformée bi-orthogonale est toujours meilleure que celle de Feauveau pour le filtrage.

- **Algorithme à trous.** Il ne crée pas d'artefact au seuillage, et les résultats sont meilleurs de manière significatives (du point de vue visuel) à haut niveau de bruit quand on le compare aux transformées orthogonales. A l'opposé des transformations (bi-) orthogonales, la nature même de l'algorithme à trous (fonction isotrope) fait qu'il est optimal quand les structures contenu dans l'image sont isotropes (comme par exemple en astronomie) et rend moins bien les contours.

- **Transformée médiane multirésolution.** Cette transformée est non-linéaire, et les estimations des niveaux de bruit dans les échelles ne peuvent se faire avec la même rigueur que dans le cas de transformations en ondelettes. Pour des bruits purement gaussiens, il n'y a clairement pas d'intérêt à utiliser une telle méthode, même si la morphologie des objets est bien respectée. Pour d'autres types de bruits, la non-linéarité peut devenir un avantage.

La transformée en ondelettes de Feauveau et la MMT ne sont pas compétitives dans le cas de bruit Gaussien. FWT-7/9 permet de mieux restaurer les contours que l'algorithme à trous, mais l'algorithme à trous est plus robuste du point de vue visuel. Un point important à noter est que la manière dont l'information est représentée est fondamentale. A haut niveau de bruit, on aura toujours plus d'artéfacts avec une transformée en ondelettes rapide qu'avec l'algorithme à trous.

Stratégies pour le traitement du bruit

- **La méthode optimale dépend du niveau de bruit.**
  A faible niveau de bruit, un simple seuillage avec une transformée en ondelettes orthogonale produit de bons résultats. Quand le bruit augmente, des artéfacts apparaissent. Des transformées non orthogonales sont alors meilleures, et les seuillages doux donnent des images d'une qualité visuelle plus acceptable.

- **L'approche de Donoho et l'approche k-sigma.**
  Quelque soit la transformée et le niveau de bruit, une approche de seuillage dur (resp. doux) à k-sigma est toujours meilleure que le seuillage dur (resp. doux) de Donoho. Aussi le PSNR que l'aspect visuel sont en faveur de l'approche k-sigma. Ceci n'est pas surprenant, car le seuil dans la méthode Donoho augmente avec le nombre de pixels (ceci étant justifié pour un avoir un nombre constant d'artéfacts). Pour une image 512×512, ça revient à seuiller à 5σ. Mais ce niveau est bien trop élevé, car la plupart des coefficients

entre 3 et 5 σ sont porteurs de beaucoup d'information. Plus l'image est grande, plus important est le sur-lissage.

- **Le seuillage hiérarchique.**
  La modification du seuil à une échelle donnée en utilisant l'information de l'échelle suivante améliore le résultat. Le PSNR est meilleur, et l'aspect visuel est similaire au seuillage dur. Cette procédure doit certainement pouvoir être appliquée dans des transformées orthogonales.

- **Entropie multiéchelle**
  L'entropie multiéchelle produit une bonne qualité visuelle quelque soit le niveau du bruit. C'est une méthode qui préserve les forts coefficients d'ondelettes, et corrige les autres coefficients d'une manière adaptative par un seuillage doux.

3.1.9 Conclusion

L'entropie multiéchelle est la meilleure méthode quand on considère l'aspect visuel et le critère PSNR. A faible niveau de bruit, une transformée en ondelettes rapide peut être utilisée, ce qui permet une meilleure restauration des contours (si l'image en contient!), tandis qu'à fort niveau de bruit, l'algorithme à trous semble plus robuste, évitant les artéfacts de la décimation. Toutefois ces artéfacts sont moins sévères que dans le cas d'un seuillage dur.

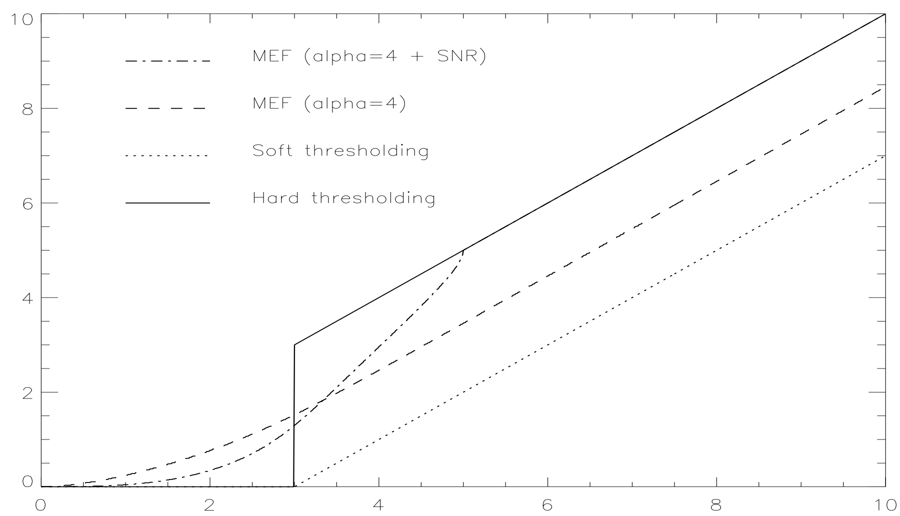

Figure 3.8: Coefficient d'ondelettes filtré en fonction du coefficient d'ondelettes (pour bruit d'écart type 1) pour quatre méthodes: seuillage dur, seuillage doux, entropie multiéchelle, entropie multiéchelle avec une valeur de α non constante (dépendante du SNR).

La figure 3.8 montre comment un coefficient d'ondelettes est modifié en utilisant un seuillage dur, un seuillage doux, l'entropie multiéchelle, et l'entropie multiéchelle avec une valeur de α non constante (dépendante du SNR). On peut voir que l'entropie multiéchelle est intermédiaire entre le seuillage dur et doux, et ne présente pas de discontinuité comme le seuillage dur. C'est pourquoi un bon SNR est obtenu avec l'entropie multiéchelle tout en gardant une bonne qualité visuelle.

3.2 Déconvolution

3.2.1 La relation objet-image

La relation objet-image est généralement donnée par:

$$I(x,y) = O(x,y) * P(x,y) + N(x,y) \tag{3.59}$$

$I(x,y)$ est l'image fournie par l'instrument de mesure, $P(x,y)$ la réponse instrumentale (PSF), $N(x,y)$ un bruit additif et $O(x,y)$ l'objet que l'on cherche à restaurer. Dans l'espace de Fourier, nous avons:

$$\hat{I}(u,v) = \hat{O}(u,v)\hat{P}(u,v) + \hat{N}(u,v) \tag{3.60}$$

On veut déterminer $O(x,y)$ connaissant $I(x,y)$ et $P(x,y)$.

Le but de la déconvolution est de trouver l'estimation la plus proche de la vraie image, compte tenu des données que l'on possède. En nommant l'ensemble des objets que l'on observe "espace objet", et l'ensemble des données dans lequel se place les images obtenues par nos capteurs (image au foyer d'un télescope ...), "espace image", on peut considérer que l'on cherche une image appartenant à l'espace objet, telle que sa projection dans l'espace image soit le plus compatible possible avec les données. L'opérateur de projection se traduit par le produit de convolution avec la réponse impulsionnelle du système d'acquisition. L'équation à résoudre correspond à un problème dit "mal posé" car il n'y a pas de solution unique et stable. Pour cette raison, les méthodes les plus performantes introduisent une régularisation (Demoment, 1989). Ceci peut être fait en introduisant dans l'opérateur inverse des connaissances a priori sur l'objet que l'on observe ou en limitant la résolution que l'on cherche à obtenir. Ces connaissances peuvent être la positivité de l'objet (contrainte de positivité), la limitation de la taille de l'objet (contrainte de support), ou des informations sur la morphologie de l'objet comme le fait l'algorithme CLEAN (Högbom, 1974) en cherchant des structures ponctuelles.

Ce problème inverse a conduit à un grand nombre de travaux. On peut les séparer en deux catégories: les méthodes non régularisées et les méthodes régularisées. Les premières sont très sensibles au bruit, tandis que les secondes sont plus robustes, mais le prix à payer est l'introduction d'une connaissance a priori sur la solution.

3.2.2 Les méthodes de déconvolution non régularisées

Filtre inverse

Cette méthode consiste à calculer la transformée de Fourier de l'objet déconvolué $\tilde{O}$ par une simple division de l'objet $\hat{O}$ par la réponse impulsionnelle $\hat{P}$.

$$\hat{\tilde{O}}(u,v) = \frac{\hat{I}(u,v)}{\hat{P}(u,v)} = \hat{O}(u,v) + \frac{\hat{N}(u,v)}{\hat{P}(u,v)} \tag{3.61}$$

Le filtre inverse nécessite de faire une transformée de Fourier et une transformée de Fourier inverse. N'étant pas itératif, il est donc très rapide. Dans les régions où la transformée de Fourier de la PSF est proche ou égale à zéro, le terme du bruit dans l'équation devient important et le bruit se trouve donc amplifié, et peut même être supérieur au signal.

Janson-Van Cittert

L'algorithme développé par Van Cittert (1931) s'écrit très simplement. En effet, on part de $k = 0$ et $O^{(k)}(x,y) = I(x,y)$ et on itère avec:

$$O^{(k+1)}(x,y) = O^{(k)}(x,y) + I(x,y) - P(x,y) * O^{(k)}(x,y) \tag{3.62}$$

quand k tend vers l'infini, on aura $O = O + I - P * O$ soit $I = P * O$. Dans l'espace de Fourier, le produit de convolution s'exprime par une multiplication, on a alors:

$$\hat{O}^{(k+1)}(u,v) = \hat{O}^{(k)}(u,v) + \hat{I}(u,v) - \hat{P}(u,v)\hat{O}^{(k)}(u,v) \tag{3.63}$$

La méthode est rapide et elle converge après seulement 5 ou 6 itérations vers la solution du filtre inverse. Par contre, la solution est instable et il est montré (Frieden, 1978a) qu'en présence de bruit, l'algorithme diverge. Jansson (1970) a modifié cette technique afin de la rendre plus résistante au bruit en introduisant un paramètre de relaxation dépendant du signal. Si on désire que $A \leq O(x,y) \leq B$, L'itération est:

$$O^{(k+1)}(x,y) = O^{(k)}(x,y) + r(x,y)[I(x,y) - P(x,y) * O^{(k)}(x,y)] \tag{3.64}$$

avec:

$$r(x,y) = C[1 - 2(B-A)^{-1} \mid O^{(k)}(x,y) - 2^{-1}(A+B) \mid]$$

C étant une constante.

Méthode de Landweber

La méthode de Landweber (1951), appelée aussi méthode du gradient à pas fixe, est en principe plus robuste que celle de Van Cittert. Elle s'écrit:

$$O^{(k+1)}(x,y) = O^{(k)}(x,y) + CP_s(x,y) * [I(x,y) - P(x,y) * O^{(k)}(x,y)] \tag{3.65}$$

avec $P_s(x,y) = P(-x,-y)$.

Ce qui est équivalent dans le plan de Fourier à:

$$\hat{O}^{(k+1)}(u,v) = \hat{O}^{(k)}(u,v) + \hat{P}_s(u,v)[\hat{I}(u,v) - \hat{P}(u,v)\hat{O}^{(k)}(u,v)] \tag{3.66}$$

avec $\hat{P}_s(u,v) = \hat{P}^*(u,v)$.

L'algorithme de Landweber minimise au sens des moindres carrés la distance entre I et $P * O$, et converge vers la solution du filtre inverse. La contrainte de positivité peut être naturellement introduite en seuillant toutes les valeurs négatives à chaque itération.

Richardson-Lucy

Développée indépendamment par W. H. Richardson (1972) et par L.B. Lucy (1974), cette technique est itérative et est déduite du théorème de Bayes sur les probabilités conditionnelles. En l'absence d'a priori sur l'objet O, l'estimation du maximum de vraisemblance donne:

$$ML(O) = \underset{O}{Max}\ p(I/O) \tag{3.67}$$

La solution vérifie donc

$$\frac{\partial \log p(I/O)}{\partial O} = 0 \tag{3.68}$$

L'hypothèse d'un bruit poissonnien conduit à:

$$[\frac{I}{P * O}] * P^* = 1 \tag{3.69}$$

En multipliant par O les deux côtés par O, l'itération de Picard (Hunt, 1994) conduit à la formule:

$$O^{(k+1)}(x,y) = [\frac{I(x,y)}{O^{(k)}(x,y) * P(x,y)} * P(-x,-y)]O^{(k)}(x,y) \tag{3.70}$$

Cette méthode est très attractive car elle conserve le flux et la positivité de l'image. De plus, elle est très simple à implanter et aussi à utiliser (aucun paramètre n'est demandé). C'est cette méthode qui a été choisie par le Space Telescope Institute pour la déconvolution des images de Hubble.

3.2.3 Méthodes régularisées

Filtre de Wiener

Le filtre de Wiener minimise au sens des moindres carrés la différence entre l'espérance mathématique de la solution et l'image restaurée. Il est caractérisé par:

$$\hat{W}(u,v) = \frac{\mid \hat{S}(u,v) \mid^2}{\mid \hat{S}(u,v) \mid^2 + \mid \hat{N}(u,v) \mid^2} \tag{3.71}$$

où $\mid \hat{S}(u,v) \mid^2$ et $\mid \hat{N}(u,v) \mid^2$ sont les densités spectrales du signal et du bruit. Le filtre déconvoluant s'écrit alors:

$$\hat{W}_d(u,v) = \frac{\hat{P}^*(u,v)}{\mid \hat{P}(u,v) \mid^2 + \frac{\mid \hat{N}(u,v) \mid^2}{\mid \hat{S}(u,v) \mid^2}} \tag{3.72}$$

Tikhonov

La régularisation de Tikhonov (1987) consiste à minimiser:

$$\parallel I(x,y) - P(x,y) * O(x,y) \parallel^2 + \lambda \parallel H(x,y) * O(x,y) \parallel^2 \tag{3.73}$$

où $H(x,y)$ est un filtre passe-haut. L'équation contient deux termes: Le premier, $\parallel I(x,y) - P(x,y) * O(x,y) \parallel^2$, exprime la fidélité par rapport aux données, et le second, $\lambda \parallel H(x,y) * O(x,y) \parallel^2$, exprime le degré de douceur dans l'image restaurée. C'est la présence de ce terme qui permet de régulariser le processus itératif. Pour cela il faut trouver la valeur optimale de λ. On peut utiliser des méthodes numériques telles que la Validation Croisée (Golub et al., 1979; Galatsanos and Katsaggelos, 1992). La méthode est longue, et produit en général des images trop lisses. Elle est inadaptée à la restauration des contours ou à la recherche de structures compactes.

Clean

CLEAN (Högbom, 1974) est la méthode la plus utilisée en radioastronomie. Elle repose sur l'hypothèse qu'un objet est composé d'un nombre discret de points sources. En prenant la transformée de Fourier inverse des données du plan de Fourier, on obtient une image appelé *dirty map* que l'algorithme CLEAN essaie de décomposer en un ensemble de pics. Ceci est fait de manière itérative en cherchant le point de plus grande intensité et en soustrayant la *dirty beam* (la *dirty beam* est obtenue par l'observation d'une source ponctuelle) multipliée par le *loop gain* (paramètre introduit dans l'algorithme) et l'intensité du pixel. L'image résultante appelée *residual map* est utilisée pour répéter le processus. La limite est atteinte quand la *residual map* est au niveau du bruit. La convolution de la liste des pics avec la *dirty beam* additionnée à la *residual map* redonne la *dirty map* . Pour obtenir l'image idéale, nommée *clean map* , il faut convoluer la liste de pics par une PSF théorique (sans lobe secondaire) et lui rajouter la *residual map*. Dans le cas d'images à faible rapport signal sur bruit, CLEAN peut trouver du signal là où il n'y a que du bruit. Pour les sources dont la taille est connue, il est possible de limiter la zone de recherche des maxima afin de diminuer le temps de calcul. La *residual map* est rajoutée afin d'avoir des images plus naturelles et de conserver la statistique du bruit. Le principe de l'algorithme est le suivant:

1. Calcul de la *dirty map* $I^{(0)}(x, y)$ et la *dirty beam* $A(x, y)$
   initialisation de la variable i à 0
2. Recherche dans $I^{(i)}(x, y)$ du point $P(x, y)$ de coordonnées (x_{max}, y_{max}) dont l'intensité I_{max} est maximale
3. Calcul de $I^{(i+1)}(x, y) = I^{(i)}(x, y) - \gamma I_{max} A_m(x, y)$
   avec $A_m(x, y) = A(x - x_{max}, y - y_{max})$
   et le *loop gain* γ compris entre 0 et 1
4. Si la *residual map* est au niveau du bruit, alors aller à 5
   sinon incrémentation de i et retour à 2
5. La *clean map* est le produit de convolution de la liste des maxima par la *clean beam* (représentée par une gaussienne)
6. L'addition de la *clean map* et de la *residual map* donne l'objet déconvolué

A chaque boucle, on retire une composante d'intensité γI_{max} au maximum de l'image, mais on retire également les lobes secondaires introduit par cette composante. La taille de la gaussienne de la *clean beam* définit la résolution que l'on veut atteindre. La gaussienne choisie est en général celle qui s'ajuste le mieux sur le lobe central de la *dirty beam*. Si on la prend plus petite, on dit qu'il y a super-résolution, car la solution contient alors de l'information à des fréquences plus grandes que la fréquence de coupure de l'instrument. Le choix de cette *clean beam* est parfois critiqué. La méthode CLEAN est un peu une "cuisine", mais elle donne de bons résultats sur les structures compactes, et est très utilisée dans le domaine de la radioastronomie car elle permet en particulier de très bien éliminer les lobes secondaires de la PSF. Dans le cas de sources étendues, où l'objet ne peut pas être assimilé à une somme de pics, les résultats sont en général médiocres. D'autres approches permettant de décomposer une image en sources ponctuelles ont été développées les dernières années (Champagnat et al., 1996; Kaaresen, 1997; Magain et al., 1998). Celle de Magain (1998) permet de plus de tenir compte d'un fond de ciel variable.

Champs de Markov

L'utilisation des Champs de Markov (Geman and Reynolds, 1992; Blanc-Féraud and Barlaud, 1996) permet de mettre une dépendance entre un pixel et ses voisins. Ceci se fait par utilisation de la distribution de Gibbs:

$$p(O) = \frac{1}{K_1} \exp(-E(O)/K_2) \tag{3.74}$$

où K_1 est une constante de normalisation, K_2 contrôle la concentration de la pdf autour de sa moyenne ($K_2 \rightarrow \infty$ signifie que la pdf est uniforme, tandis $K_2 \rightarrow 0$ conduit à une pdf concentrée autour d'un mode spécifique associé à la fonction $E(O)$. $E(O)$ est la fonction d'énergie qui prend la forme:

$$E(O) = \sum_{c \in \mathcal{C}} U_c(O) \tag{3.75}$$

où $\mathcal{C}$ est un ensemble de cliques c correspondant au voisinage $\mathcal{V}$ et $U_c(O)$ sont les *fonctions de potentiel d'interaction locale* qui définissent des *contraintes locales* sur le champ O à estimer. Ces fonctions attribuent un coût (i.e. *énergie locale*) à chaque configuration spéciale de pixels c, appelé *clique*. Les cliques d'ordre deux sont définis par les couples de pixels voisins (s,t). Un exemple de fonction potentiel est:

$$U(s,t) = (O(s) - O(t))^2 \tag{3.76}$$

Dans le cas Gaussien, on déduit alors de l'équation de Bayes que la fonctionnelle à minimiser est

$$J(O) = \frac{\parallel I - P * O \parallel}{2\sigma^2} + \lambda \sum_{c \in \mathcal{C}} U_c(O) \tag{3.77}$$

où λ est le paramètre de régularisation.

Maximum d'entropie

La méthode MEM (Frieden, 1978a) (*Maximum Entropy Method*), proposée par le radioastronome Ables, est une approche probabiliste du problème. La méthode consiste à choisir comme solution la plus probable celle qui maximise l'entropie. L'entropie caractérisant le taux d'information contenu dans l'image, MEM recherche l'image la plus simple possible compatible avec les données et les connaissances a priori sur l'objet. Il y a une controverse sur la définition de l'entropie (voir la discussion au premier chapitre). La plus utilisée est sans doute l'entropie de Gull et Skilling (1991).

La multirésolution a été un des principaux moteurs pour le développement de nouvelles méthodes au cours des dernières années. La nécessité d'introduire ce concept s'est fait sentir dès 1988 avec l'algorithme Muliresolution Clean. Les travaux de Weir (Weir, 1992) et (Bontekoe et al., 1994) vont aussi dans cette direction. Puetter (Puetter, 1986) introduit un dictionnaire de formes multirésolutions. Dans ce cas, la connaissance a priori que l'on doit avoir est très forte. Plus récemment, Nunez et Llacer (Nunez and Llacer, 1998) ont proposé d'introduire un pas en intensité variable, suivant les zones de l'image, ces zones étant définies interactivement par l'utilisateur à partir de la transformée en ondelettes de l'image.

3.2.4 Ondelettes et déconvolution

La transformation en ondelettes a été utilisée de différentes manières pour développer de nouvelles approches de régularisation. La plupart des méthodes peuvent en fait être "repensées" dans un cadre ondelette. Nous présentons ici un éventail de possibilités.

Clean Multi-Resolution

Pour pallier le problème que posait CLEAN sur les images à grandes structures, Wakker et Schwarz (1988) ont proposé un nouvel algorithme reposant sur une analyse en multirésolution (MRC). La méthode consiste à appliquer CLEAN non plus sur la *dirty map* , mais sur deux images appelées *smooth dirty map* et *difference dirty map* . La *smooth dirty map* est obtenue par lissage de la *dirty map* tandis que la *difference dirty map* est la différence entre la *dirty map* et la *smooth dirty map* . En effectuant les mêmes opérations sur la *dirty beam* et la *clean beam* , on obtient la *smooth dirty beam* , la *difference dirty beam* , la *smooth clean beam* et la *difference clean beam* . On pose:

- G = la fonction de lissage normalisée ($\int G(x)dx = 1$)
- A = *dirty beam*
- D = *dirty map*
- δ = fonction-δ
- R = *residual map* après le passage de CLEAN sur la MAP
- B = *clean beam* normalisée
- C = *clean map*
- s = facteur de normalisation de la *dirty beam*
- r = facteur de normalisation de la *clean beam*
- A_s = *smooth dirty beam* normalisée = $sA * G$
- A_d = *difference dirty beam* normalisée = $\frac{1}{1-\frac{1}{s}}(A - \frac{A_s}{s})$
- B_s = *smooth clean beam* normalisée = $rB * G$
- B_d = *difference clean beam* normalisée = $\frac{1}{1-\frac{1}{r}}(B - \frac{B_s}{r})$

A partir des pics trouvés par CLEAN, on peut retrouver la *dirty map* en les convoluant par la *dirty beam* , et en rajoutant la *residual map.*

$$D = D_s + D_d = \delta_s * A_s + R_s + \delta_d * A_d + R_d \tag{3.78}$$

ce qui peut aussi s'écrire:

$$D = [s\delta_s * G + \frac{s}{s-1}\delta_d * (1 - G)] * A + R_s + R_d \tag{3.79}$$

Si on remplace la *dirty beam* par la *clean beam* , on obtient la *clean map* :

$$C = \frac{s}{r}\delta_s * B_s + \frac{s(r-1)}{r(s-1)}\delta_d * B_d + R_s + R_d \tag{3.80}$$

L'algorithme *Multi-Resolution Clean* (MRC) nécessite trois paramètres. Le premier fixe la fonction de lissage G, et les deux autres sont le *loop gain* et l'*extra loop gain* utilisés par CLEAN sur la *smooth dirty map* et la *difference dirty map*. Il a été montré que cette méthode peut se décrire dans un cadre ondelette, et peut être améliorée en utilisant une reconstruction itérative (Starck et al., 1994; Starck and Bijaoui, 1994b).

Régularisation à partir des structures significatives.

On note $R^{(n)}(x,y)$ l'erreur à l'itération n d'un algorithme de déconvolution, on a:

$$R^{(n)}(x,y) = I(x,y) - I^{(n)}(x,y) \tag{3.81}$$

Si l'approximation $O^{(n)}(x,y)$ est très proche de l'objet $O(x,y)$, $R^{(n)}(x,y)$ doit correspondre au bruit $N(x,y)$. Au lieu d'introduire une contrainte de régularisation du type Tikhonov, ou le Maximum d'Entropie (Frieden, 1978a), on peut simplement chercher à examiner si le résidu peut être ou non considéré comme un bruit (Starck and Murtagh, 1994). Seules les structures significatives du résidu sont alors prises en compte. Notons $\bar{R}^{(n)}$ le résidu significatif, les itérations des méthodes régularisées de Van Cittert, Landweber et Lucy se font de la manière suivante (Starck and Murtagh, 1994):

- Régularisation de l'algorithme de Van Cittert
  L'itération de Van Cittert est:

  $$O^{(n+1)}(x,y) = O^{(n)}(x,y) + \alpha R^{(n)}(x,y) \tag{3.82}$$

  avec $R^{(n)}(x,y) = I(x,y) - P(x,y) * O^{(n)}(x,y)$. La régularisation par les structures significatives conduit à:

  $$O^{(n+1)}(x,y) = O^{(n)}(x,y) + \alpha \bar{R}^{(n)}(x,y) \tag{3.83}$$

  Le principe consiste donc à détecter à chaque échelle, les structures significatives dans le résidu $R^{(n)}(x,y)$, et à les transférer dans l'objet restauré $O^{(n)}(x,y)$. Le processus se finit quand on ne détecte plus rien. Alors, on a séparé $I(x,y)$ en deux images $\tilde{O}(x,y)$ et $R(x,y)$. $\tilde{O}$ est l'image restaurée qui ne contient pas de bruit, et $R(x,y)$ est le résidu final qui ne contient pas de structure. $R(x,y)$ est en fait notre estimation du bruit $N(x,y)$.

- Régularisation de la méthode du gradient à pas fixe
  L'itération de la méthode du gradient à pas fixe est:

  $$O^{(n+1)}(x,y) = O^{(n)}(x,y) + \alpha P(-x,-y) * R^{(n)}(x,y) \tag{3.84}$$

La régularisation conduit à:

$$O^{(n+1)}(x,y) = O^{(n)}(x,y) + \alpha P(-x,-y) * \bar{R}^{(n)}(x,y) \tag{3.85}$$

- Régularisation de l'algorithme de Lucy

  Notons $I^{(n)}(x,y) = P(x,y) * O^{(n)}(x,y)$, soit $I(x,y) = I^{(n)}(x,y) + R^{(n)}(x,y)$. L'équation de Lucy peut être récrite sous la forme suivante:

$$O^{(n+1)}(x,y) = O^{(n)}(x,y)[\frac{I^{(n)}(x,y) + R^{(n)}(x,y)}{I^{(n)}(x,y)} * P(-x,-y)] \tag{3.86}$$

  La régularisation conduit à:

$$O^{(n+1)}(x,y) = O^{(n)}(x,y)[\frac{I^{(n)}(x,y) + \bar{R}^{(n)}(x,y)}{I^{(n)}(x,y)} * P(-x,-y)] \tag{3.87}$$

L'écart type du résidu diminue jusqu'à ce qu'il n'y ait plus de structures significatives détectées. La convergence peut être estimée à partir de l'écart type du résidu. L'algorithme est arrêté quand:

$$\frac{\sigma_{R^{(n-1)}} - \sigma_{R^{(n)}}}{\sigma_{R^{(n)}}} < \epsilon \tag{3.88}$$

Régularisation à partir du Support Multirésolution

On peut ne considérer que la position du coefficient multirésolution plutôt que son amplitude. $\bar{R}^{(n)}$ représentera alors le résidu contenu dans le support multirésolution (Murtagh et al., 1995b; Starck et al., 1995). Cette manière de procéder permet d'avoir une meilleure photométrie dans le cas où l'image contient des étoiles. La méthode consistera donc à calculer le support multirésolution à partir des données, puis à chaque itération du processus de déconvolution, à ne tenir compte que de l'information contenue dans ce support.

L'utilisation des structures significatives, ou du support multirésolution permet de régulariser un processus de déconvolution. L'estimation du support multirésolution étant faite de manière automatique, cette méthode est très simple à mettre en oeuvre avec une bonne efficacité.

3.2.5 Principe de la déconvolution par entropie multiéchelle

Introduction

Le principe de la méthode MEM a été décrit précédemment. La solution la plus réaliste est celle qui minimise le montant d'information, tout en restant compatible avec les données.

$$J(O) = \sum_{k=1}^{N} \frac{(I_k - (P * O)_k)^2}{2\sigma_I^2} + \alpha H(O) \tag{3.89}$$

Nous avons vu que dans le cas Gaussien, H est donné pas l'énergie des coefficients d'ondelettes, et ses propriétés sont proches de celle de l'entropie de Shannon. On a alors

$$J(O) = \sum_{k=1}^{N} \frac{(I_k - (P * O)_k)^2}{2\sigma_I^2} + \alpha \sum_{j=1}^{l} \sum_{k=1}^{N_j} \frac{w_{j,k}^2}{2\sigma_j^2} \tag{3.90}$$

où σ_j est l'écart type du bruit à l'échelle j, N_j le nombre de pixels à l'échelle j, σ_I l'écart type du bruit dans les données, et l le nombre échelles.

Plutôt que de minimiser le montant d'information dans la solution, on préfère minimiser le montant d'information qui peut être dû au bruit. La fonctionnelle devient

$$J(O) = \sum_{k=1}^{N} \frac{(I_k - (P * O)_k)^2}{2\sigma_I^2} + \alpha H_n(O) \tag{3.91}$$

En utilisant l'approche N2-MSE, H_n est défini par

$$H_n(w_{j,k}) = \sum_{j=1}^{l} \sum_{k=1}^{N_j} \int_0^{|w_{j,k}|} P_n(\mid w_{j,k} \mid -u)(\frac{\partial h(x)}{\partial x})_{x=u} du \tag{3.92}$$

ce qui donne pour un bruit Gaussien

$$H_n(X) = \sum_{j=1}^{l} \sum_{k=1}^{N_j} \frac{1}{\sigma_j^2} \int_0^{|w_{j,k}|} u \ \mathrm{erf}(\frac{\mid w_{j,k} \mid -u}{\sqrt{2}\sigma_j}) \tag{3.93}$$

La solution est trouvée en calculant le gradient $\nabla(J(O))$, et en effectuant l'itération suivante

$$O^{n+1} = O^n - \gamma \nabla(J(O^n)) \tag{3.94}$$

où γ est un paramètre de convergence.

Comme pour le filtrage, il y a plusieurs façons d'utiliser le paramètre α qui fixe le degré de douceur. On peut considérer un α_j par échelle, et introduire une sorte de régularisation adaptative, dépendante du rapport signal sur bruit des coefficients d'ondelettes des données.

Les paramètres

Afin de laisser une certaine flexibilité dans la manière de restaurer les données, on introduit deux paramètres $\beta_{j,k}$ et $\alpha_{j,k}$ qui nous permettent de pondérer respectivement les deux termes de l'équation à minimiser:

$$J(O) = \frac{1}{2\sigma_I^2} \sum_{k=1}^{N} (c_{l,k}(R) + \sum_{j=1}^{l} \beta_{j,k} w_{j,k}(R))^2 + \sum_{j=1}^{l} \sum_{k=1}^{N_j} \alpha_{j,k} h(w_{j,k}(O)) \tag{3.95}$$

où $R = I - P * O$, et

$$R = c_{l,k}(R) + \sum_{j=1}^{l} w_{j,k}(R)$$

$w_{j,k}(R)$ sont les coefficients en ondelettes de R, et $w_{j,k}(O)$ sont les coefficients en ondelettes de O.

Le paramètre $\beta_{j,k}$

On considère trois approches pour estimer $\beta_{j,k}$

1. pas de pondération: $\beta_{j,k} = 1$
2. pondération douce: $\beta_{j,k} = P_s(w_{j,k}(I))$
   Dans ce cas, $\beta_{j,k}$ est égal à la probabilité que les coefficients en ondelettes des données soient du signal (et pas du bruit).
3. pondération dure: $\beta_{j,k} = 0$ ou 1 en fonction $P_n(w_{j,k}(I))$ (avec $P_n(w_{j,k}(I)) = 1 - P_s(w_{j,k}(I))$).
   Ça consiste à n'utiliser que les coefficients significatifs des données. Si M est le support multirésolution de l'image d'entrée I, alors $\beta_{j,k} = M(j,k)$.

Le paramètre $\alpha_{j,k}$

Le paramètre $\alpha_{j,k}$ est le produit de deux valeurs: $\alpha_{j,k} = \alpha_u \beta'_{j,k}$.

- α_u est un paramètre utilisateur (1 par défaut) qui permet de contrôler la douceur de la solution. Augmenter α_u produit une solution plus lisse.
- $\beta'_{j,k}$ dépend des données d'entrée et peut être défini de la manière suivante
  1. pas de régularisation ($\beta'_{j,k} = 0$): Seul le premier terme de la fonctionnelle est minimisé.
  2. pas de protection contre la régularisation ($\beta'_{j,k} = 1$): la régularisation est effectuée à toutes les positions et à toutes les échelles.
  3. protection douce ($\beta'_{j,k} = P_n(w_{j,k}(I))$): la régularisation devient adaptative, dépendante du rapport signal sur bruit des coefficients d'ondelettes des données d'entrée.
  4. protection dure($\beta'_{j,k} = 0$ ou 1, en fonction de $P_n(w_{j,k}(I))$). Si M est le support multirésolution de l'image d'entrée I, alors $\beta_{j,k} = 1 - M(j,k)$.
  5. protection douce + dure: ($\beta'_{j,k} = 0$ ou $P_n(w_{j,k}(I))$ en fonction de $P_n(w_{j,k}(I))$).

Notons que choisir une pondération dure et pas de régularisation conduit à une déconvolution par la méthode du support multirésolution.

La fonction de corrélation intrinsèque (ICF)

Dans beaucoup de cas, déconvoluer une image jusqu'à atteindre la résolution du pixel (surtout si la PSF est extrêmement large) n'a pas réellement de sens. L'idée de limiter la résolution que l'on veut atteindre est relativement ancienne, car elle est déjà contenue dans le concept de la méthode CLEAN (Högbom, 1974). En effet, la *Clean-Beam* fixe la résolution de la solution finale. Ce principe a été aussi développé par Lannes (Lannes and Roques, 1987) d'une manière plus formelle. Depuis, le principe est de temps en temps réinventé, Gull et Skilling d'abord, la *Clean-Beam* étant appelée "*Intrinsic Correlation Function*" (ICF), puis plus récemment Magain et al (Magain et al., 1998). L'ICF est normalement une Gaussienne, mais dans certains cas, il peut être intéressant de prendre une autre fonction. Par exemple, si on veut comparer deux images I_1 et I_2, obtenues à des longueurs d'onde différentes, la méthode classique est de convoluer I_1 par P_2 et I_2 par P_1 (P_1 et P_2 étant respectivement les PSF de I_1 et I_2). Les images sont alors à la même résolution. Mais malheureusement, il peut y avoir une grosse perte de résolution en effectuant ces convolutions. Déconvoluer les deux images n'est en général pas possible, car il est impossible de connaître la résolution atteinte, qui n'est pas forcément la même pour les deux déconvolutions. La solution est de ne déconvoluer que l'image qui a la plus mauvaise résolution (disons I_1), et de limiter la déconvolution à la résolution de I_2. Pour cela, il suffit de prendre P_2 comme ICF. Le problème de cette déconvolution est donc de trouver $\tilde{O}$ (solution cachée) tel que:

$$I_1 = P_1 * P_2 * \tilde{O} \tag{3.96}$$

et la solution O_1, à la même résolution que I_2, est obtenue par une simple convolution de $\tilde{O}$ par P_2. Les images O_1 et I_2 peuvent ensuite être comparées.

Introduire une ICF G dans la déconvolution multiéchelle revient à considérer une nouvelle PSF P' qui est la convolution de P et G. La déconvolution est faite en utilisant P', et la solution doit être convoluée à la fin par G .

Exemple: déconvolution de Beta Pictoris

L'image Beta Pictoris (Pantin and Starck, 1996) a été obtenue avec la caméra infrarouge TIMMI, placée sur le télescope de 3m60 de l'ESO. La forme du disque a été partiellement perdue à cause de l'effet de la réponse instrumentale. La déconvolution s'impose alors afin d'obtenir le profil exact du disque, et de pouvoir comparer le résultat aux modèles théoriques.

Après filtrage (voir figure 3.10 à gauche), le disque apparaît clairement. L'image déconvoluée (figure 3.10 à droite) montre que le disque est étendu et asymétrique. L'entropie multiéchelle s'est révélée plus appropriée pour la régularisation d'une telle image que les autres méthodes, et permet une bonne reconstruction des structures les plus faibles du disque.

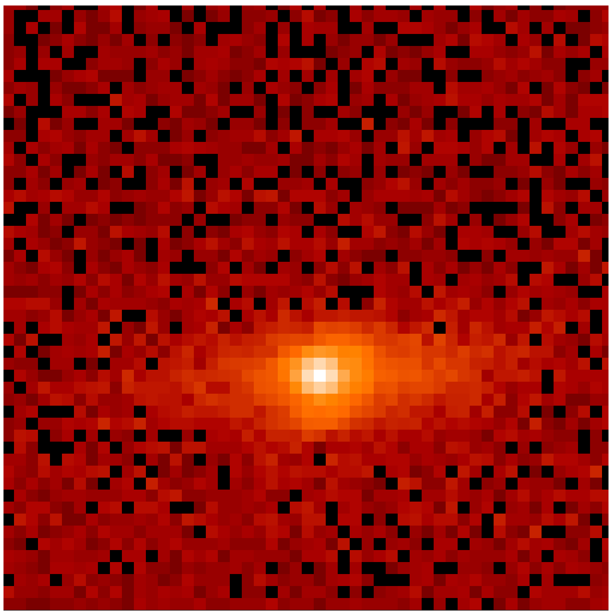

Figure 3.9: image TIMMI de Beta Pictoris.

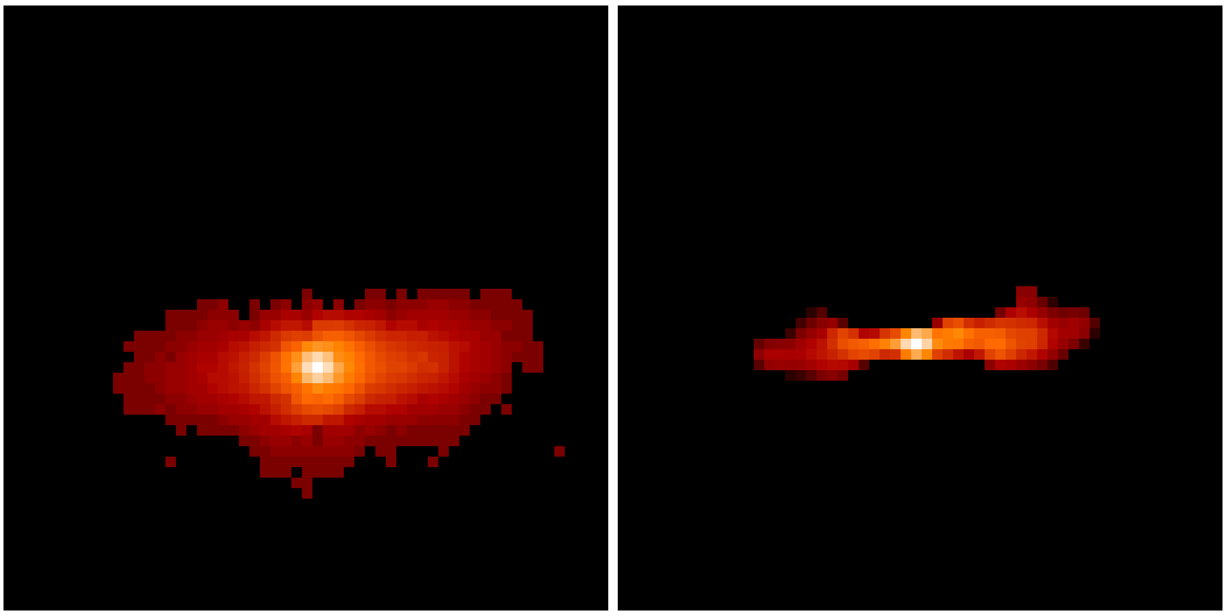

Figure 3.10: Image filtrée (gauche), et image déconvoluée (droite).

3.3 Analyse de fluctuation du fond du ciel

L'entropie multiéchelle a été définie par

$$H(X) = -\sum_{j=1}^{l}\sum_{k=1}^{N} h(w_{j,k}) \tag{3.97}$$

où $h(w_{j,k}) = \ln(p(w_{j,k})))$.

Afin d'étudier le comportement du bruit à travers les échelles, on définit le vecteur d'entropie moyen E par (Gastaud et al., 1999)

$$E(j) = \frac{1}{N}\sum_{k=1}^{N} h(w_{j,k}) \tag{3.98}$$

$E(j)$ donne la valeur de l'entropie moyenne à l'échelle j. A partir du vecteur E, on a une information statistique sur chaque échelle séparément. Étant donné un modèle de bruit, on est capable de calculer (en général à partir de simulations) le vecteur d'entropie moyen lié au bruit $E^{(bruit)}(j)$, c'est à dire la mesure d'information dans le cas ou les données ne contiennent que du bruit. On préfère alors le vecteur normalisé

$$E_n(j) = \frac{E(j)}{E^{(bruit)}(j)} \tag{3.99}$$

La figure 3.11 montre le résultat d'une simulation. Cinq images ont été créées en ajoutant n sources à une image 1024x1024 contenant un bruit Gaussien d'écart type 1. Toutes les sources sont identiques, ce sont des Gaussiennes avec un maximum égal à 1, et d'écart type égal à 2. En définissant le rapport signal sur bruit (SNR) comme le rapport entre l'écart type de la source dans une boite contenant 90 % du flux et l'écart type du bruit, on a un SNR de 0.25. Les sources sont absolument indétectables dans les image simulées, aussi bien que dans leurs transformées en ondelettes. La figure 3.12 montre une région qui contient une source au centre. Il est évidemment impossible de trouver une telle source noyée dans le bruit. Les cinq images ont été obtenues en utilisant un nombre de sources respectivement égal à 0,50,100,200 et 400. La simulation a été répétée dix fois avec différentes cartes de bruit afin d'obtenir des barres d'erreur sur les mesures d'entropie. Le nombre de pixels affectés par les sources dans les images est inférieur à 2.5%.

Quand le nombre de sources augmente, la différence entre les courbes, représentant le vecteur d'entropie moyen normalisé, augmente. Même si les sources sont très faibles, la présence de signal peut être mise en évidence. Par contre, la position de ces sources reste inconnue.

L'entropie multiéchelle permet donc de mettre en évidence une variation dans le fond du ciel qui ne peut pas être liée au modèle de bruit considéré. Cette approche peut être une alternative aux méthodes actuellement utilisées comme la fonction de corrélation à deux points.

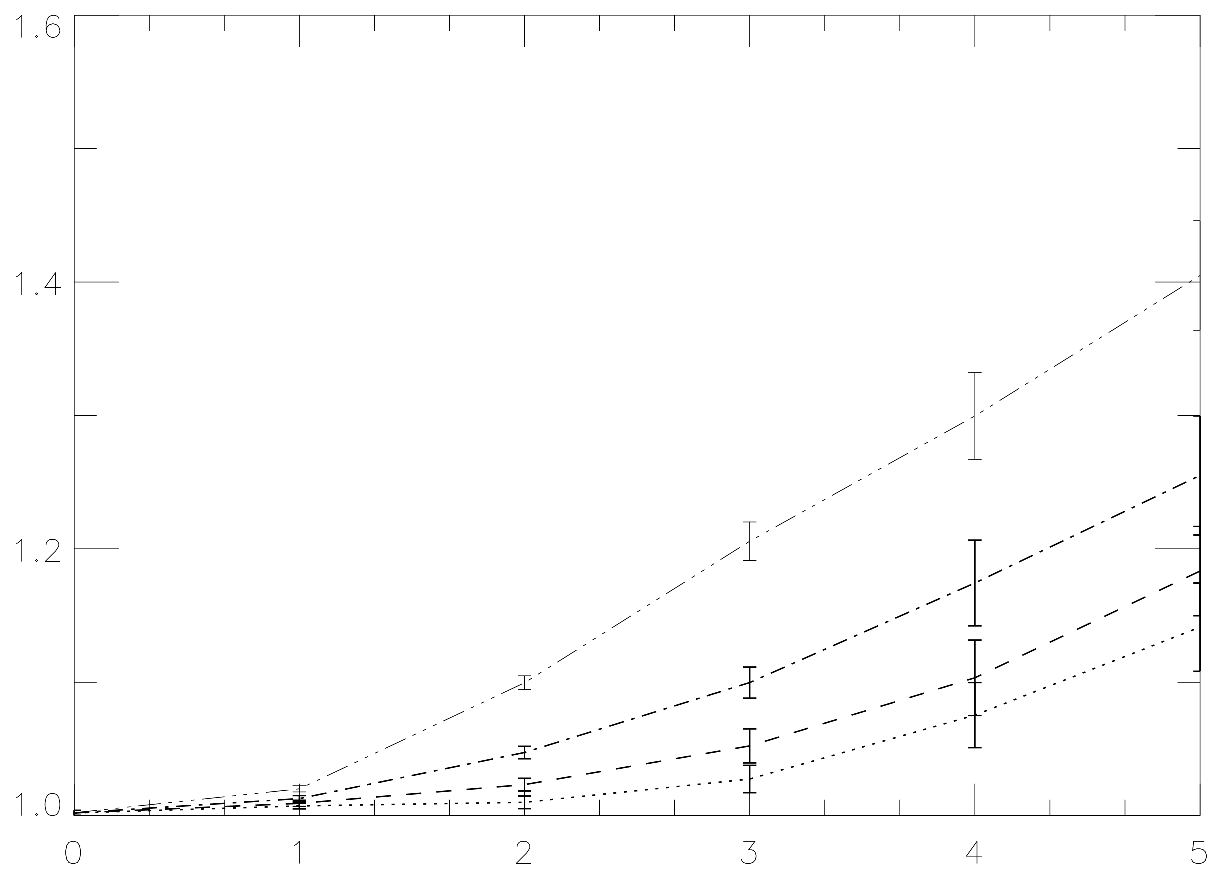

Figure 3.11: Entropie moyenne en fonction de l'échelle pour cinq images simulées contenant des sources noyées dans du bruit. Chaque courbe correspond au vecteur d'entropie moyen d'une image. De haut en bas, l'image contient respectivement 400,200,100, 50 et 0 sources.

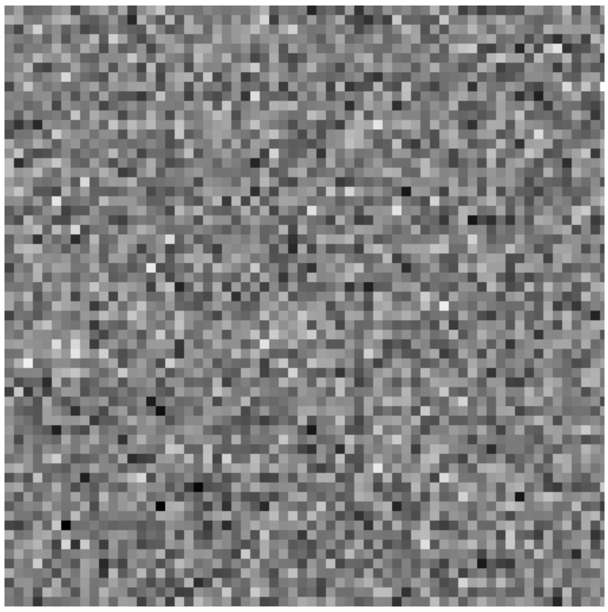

Figure 3.12: Région d'une image simulée contenant une source faible au centre.

3.4 Conclusions générales

L'entropie multiéchelle est une nouvelle manière de mesurer l'information contenue dans un signal. Chaque échelle est porteuse d'une certaine quantité d'information, et la somme de ces quantités donne l'information globale. Cette approche est en fait utilisée depuis de nombreuses années dans le domaine de la compression d'images. En effet, les méthodes modernes de compression de données consistent à appliquer d'abord une transformation (transformée cosinus, transformée ondelettes, ...), puis à coder les coefficients obtenus. Une bonne transformée pour la compression de données est en général une transformée orthogonale, car il n'y a pas de redondance de l'information. L'entropie de Shannon, appliquée à l'image transformée, donne alors le nombre exact de bits nécessaires au codage de ces coefficients. Pour la restauration de signal, le problème n'est pas de réduire le nombre de coefficients dans la représentation multiéchelle des données, et on préfère en général utiliser une transformée non orthogonale qui évite les artefacts lors de la reconstruction.

On aurait pu utiliser l'entropie de Shannon pour mesurer l'information à une échelle donnée, mais il nous a semblé nécessaire d'introduire la loi de probabilité du bruit dans notre mesure. Nous avons vu qu'une telle approche conduit à une relation très physique entre l'information et les coefficients d'ondelettes dans le cas d'un bruit gaussien: l'information est proportionnelle à l'énergie des coefficients d'ondelettes normalisés par l'écart type du bruit. Cette approche peut être généralisée à d'autres types de bruit.

Pour le filtrage, l'entropie multiéchelle présente les avantages suivant:

- C'est un bon compromis entre seuillage dur et mou.
- Aucun modèle a priori sur le signal lui même n'est nécessaire, ce qui n'est pas le cas avec les autres méthodes bayesiennes proposées dans la littérature (Chipman et al., 1997; Crouse et al., 1998; Vidakovic, 1998; Timmermann and Nowak, 1999).
- Elle peut être généralisée à d'autres classes de bruit.
- Le paramètre α de régularisation peut être fixé automatiquement.

Nous avons aussi vu que remplacer la mesure d'entropie standart par l'entropie multiéchelle résout les principaux problèmes relatifs à la méthode de déconvolution MEM. Finalement, l'entropie multiéchelle, qui a été conçue dans un objectif de restauration de signal s'est avérée être également utile pour l'analyse de données, comme l'étude de la fluctuation de fond du ciel.

Les travaux menés actuellement en collaboration avec le Centre de Données astronomiques de Strasbourg (CDS) montrent la pertinence de l'entropie multiéchelle (Starck et al., 1999c). En effet, une relation quasi linéaire est obtenue entre l'entropie multiéchelle et le taux de compression optimal d'une image astronomique (voir figure 3.13). Le taux de compression est optimal pour

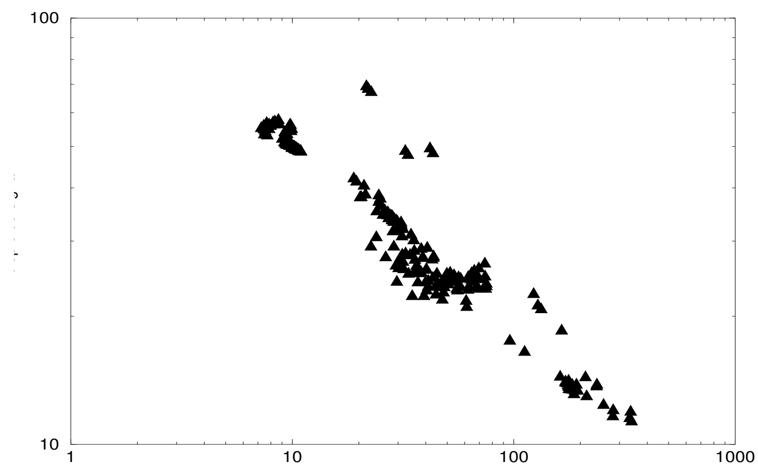

Figure 3.13: Entropie multiéchelle d'images astronomiques en fonction du taux optimal de compression. Les images qui contiennent un grand nombre de sources ont un petit taux de compression et une grande entropie multiéchelle. La relation est presque linéaire.

un critère donnée, et dans le cas d'images astronomiques, le critère est l'erreur acceptable sur les mesures astrométriques et photométriques des différents objets contenus dans l'image. L'entropie multiéchelle est donc un prédicteur du taux optimal de compression que l'on peut attendre sur une image astronomique. Dans le cadre de la compression de grandes images, cette propriété de linéarité devrait permettre de fixer la taille des sous-images à comprimer afin d'avoir des fichiers de taille fixe accessibles aux utilisateurs par le réseau.

D'autres applications sont aussi envisageables dans le futur comme la recherche d'information dans les bases de données d'images astronomiques. En effet, Csillaghy (1996) a montré l'utilité d'informations extraites des images pour optimiser les recherches et l'entropie multiéchelle peut certainement jouer un rôle au cours de cette phase.

Chapitre 4

Travaux, Encadrement et Enseignement

4.1 Travaux et Projets

4.1.1 Thèse (1990-1992)

J'ai effectué ma thèse sous la direction d'Albert Bijaoui à l'observatoire de Nice et de Gérard Schumacher au CERGA. Le sujet était:
Analyse en Ondelettes et Imagerie à Haute Résolution Angulaire.
Le but de cette thèse était d'étudier comment l'analyse en ondelettes pouvait être utilisée pour la reconstruction d'images interférométriques.

Dans un premier temps, les problèmes liés à la reconstruction d'images à haute résolution angulaire par voie interférométrique dans le visible ont été analysés. Pour ce faire, il a été nécessaire de développer une chaîne de simulation comportant:

- Un simulateur de données interférométriques, qui génère des séries d'interférogrammes dispersés en longueur d'onde, en comptage de photons et perturbés par une turbulence de type atmosphérique introduisant des phases aléatoires.

- Des algorithmes de restauration de la phase (analyse bispectrale, calibration par espacements redondants, ...).

- Des algorithmes de déconvolution (Clean, Lucy, ...).

Les résultats de ces simulations ont été publiés dans JOSA A (P. Cruzalèbes and Starck, 1992).

Ensuite, une étude approfondie de l'analyse en ondelettes a montré les difficultés que présentait la non-isotropie de l'ondelette de Mallat et nous a amené à créer un algorithme de transformation en ondelettes reposant sur des fonctions à fréquence de coupure complètement isotropes.

Cet algorithme permet d'associer à une image une structure pyramidale, chaque plan de la pyramide correspondant aux détails de l'image à une échelle donnée. Différentes applications ont ensuite été déduites de cette pyramide (Starck and Bijaoui, 1994a):

- Filtrage: Plusieurs méthodes de filtrage ont été mises au point. L'utilisation de l'information hiérarchique de la matière, c'est à dire les dépendances inter-échelles, pour élaborer le filtre, a conduit à un filtre de Wiener hiérarchique.
- Critère Qualité: Les critères de qualité habituels (corrélation, entropie, écart quadratique ...) ne tiennent pas compte de la résolution. La comparaison de deux images dans l'espace des ondelettes échelle par échelle donne ainsi une information supplémentaire qui est utile pour une évaluation objective des méthodes.
- Déconvolution: En déduisant l'ondelette de la réponse instrumentale, une nouvelle approche de la déconvolution a été proposée.

La dernière étape était donc de faire le lien entre l'analyse en ondelettes et la déconvolution de données interférométriques. Ce travail a été accompli en étudiant dans un premier temps le problème de la déconvolution d'une manière générale dans l'espace des ondelettes (Starck and Bijaoui, 1994b). Dans un deuxième temps, une méthode de déconvolution pour la synthèse d'ouverture a été déduite en généralisant dans l'espace des ondelettes la méthode de Wakker (1988) et en introduisant des contraintes de régularisation. Le principe de la méthode est de détecter les structures significatives, que l'on pourrait appeler atomes fondamentaux, à chaque échelle. Cela revient en fait à rechercher une représentation de l'image avec le minimum d'atomes. La reconstruction se fait ensuite itérativement à partir de ces atomes. La régularisation est due au fait que les atomes sont moins nombreux que les visibilités mesurées et que d'un problème mal-posé, on a su se placer dans le cadre d'un problème bien posé. La position de ces atomes dans l'espace des ondelettes impose en fait une contrainte de support multirésolution au cours du processus de restauration. Cette contrainte, extrêmement forte, permet de restaurer une image sans aucun artefact.

Une collaboration avec l'équipe Soir d'été de J. Gay a permis de reconstruire les étoiles à enveloppe Hen 1379 et OH 26.5+0.6 dont les données ont été obtenues par speckle interférométrie dans le proche infrarouge (Starck et al., 1994). La figure 4.2 montre la reconstruction de Hen 26.5+0.6. Les ondelettes ont permis de mettre en évidence l'enveloppe de l'étoile, ainsi que son assymétrie.

4.1.2 Un contexte Ondelette dans MIDAS

Au cours de mon séjour (en 1993) à l'ESO (*European Southern Observatory*), j'ai créé un contexte ondelette dans le logiciel d'analyse de données MIDAS (version NOV93) (Starck, 1993; Murtagh et al., 1993). Ce contexte inclut un programme de démonstration des différents algorithmes de la transformée en ondelettes, le type de données qu'elles produisent, et la façon de les visualiser. Plusieurs algorithmes et outils de manipulation de la transformée en ondelettes ont été implantés (23 commandes au total), ainsi que des programmes de filtrage et de déconvolution.

4.1.3 Déconvolution des images du Hubble Space Telescope

En collaboration avec le groupe du ST-ECF, j'ai travaillé sur le problème de la déconvolution des données du Hubble Space Telescope (Murtagh et al., 1995a) (problème crucial étant donné que le HST n'était pas réparé à cette époque). Ces travaux nous ont amené a définir le concept de support multirésolution (Starck and Murtagh, 1994; Murtagh et al., 1995a; Murtagh et al., 1995b; Starck et al., 1995; Bijaoui et al., 1994) d'une image, et de l'utiliser comme moyen de régularisation afin d'empêcher l'amplification du bruit au cours du processus itératif de déconvolution.

Le support multirésolution d'une image $I(x,y)$ est défini de la manière suivante:

$$M(j,x,y) = \begin{cases} 1 & \text{si } w_j(x,y) \text{ est significatif} \\ 0 & \text{si } w_j(x,y) \text{ n'est pas significatif} \end{cases} \tag{4.1}$$

où j est l'échelle.

L'algorithme permettant de générer le support multirésolution est le suivant:

1. On calcule la transformée en ondelettes de l'image.
2. On estime le niveau du bruit à chaque échelle. On en déduit statistiquement le seuil de signification.
3. La binarisation de chaque échelle de la transformée en ondelettes conduit au support multirésolution.
4. On modifie le support dans le cas où l'on a une connaissance a priori du contenu de l'image.

La dernière étape est évidemment optionnelle. Un exemple type où l'introduction d'une telle connaissance est utile est le cas où la réponse instrumentale est relativement étalée. En effet, on peut alors supprimer dans le support tous les pixels isolés qui sont certainement dus à du bruit.

L'algorithme de déconvolution de Lucy (1974) est

$$\begin{aligned} O^{(n+1)} &= O^{(n)}[(I/I^{(n)}) * P^*] \\ I^{(n)} &= P * O^{(n)} \end{aligned} \tag{4.2}$$

où O est l'objet à observé (inconnu), I les données et P la réponse instrumentale. Cette équation peut aussi s'écrire sous la forme:

$$O^{(n+1)}(x,y) = O^{(n)}(x,y)\left[\frac{I^{(n)}(x,y) + R^{(n)}(x,y)}{I^{(n)}(x,y)} * P(-x,-y)\right] \tag{4.3}$$

où $R^{(n)}$ est le résidu à l'itération n ($R^{(n)} = I - P * O^{(n)}$). On décompose alors le résidu R par l'algorithme à trous

$$R^{(n)}(x,y) = c_p(x,y) + \sum_{j=1}^{p} w_j(x,y) \tag{4.4}$$

puis on ne tient compte que de ce qui est contenu dans le support multirésolution

$$\bar{R}^{(n)}(x,y) = c_p(x,y) + \sum_{j=1}^{p} M(j,x,y)\ w_j(x,y) \tag{4.5}$$

L'algorithme de Lucy régularisé par les ondelettes devient alors:

$$O^{(n+1)}(x,y) = O^{(n)}(x,y) \left[\frac{I^{(n)}(x,y) + \bar{R}^{(n)}(x,y)}{I^{(n)}(x,y)} * P(-x,-y) \right] \tag{4.6}$$

L'algorithme de Van Cittert ou du gradient peut être régularisé de la même manière.

La figure 4.3 représente une simulation d'image du HST (*Wide Field Camera*) d'un amas distant de galaxies. L'image a été convoluée par la réponse instrumentale du HST et un bruit a été ajouté. Les images du bas montrent les restaurations par la méthode de Lucy et par la méthode de Lucy régularisée. Des techniques de détection automatique d'objets astronomiques ont ensuite été utilisées sur ces deux images. Les objets trouvés ont pu être alors comparés à ceux existant réellement dans l'image de départ (en haut à gauche). Les résultats de cette étude montrent que:

- La suppression du bruit aide fortement la détection d'objets. Un nombre plus important d'objets réels a été détecté avec la méthode de Lucy régularisée.
- La précision des coordonnées des centroides des objets est plus élevé avec la méthode de Lucy régularisée.
- Le rapport signal sur bruit de l'image dégradée est de 3.5 dB, celui de l'image restaurée par la méthode de Lucy est de 10dB, et celui de l'image restaurée par la méthode de Lucy régularisée est de 15.5 dB.

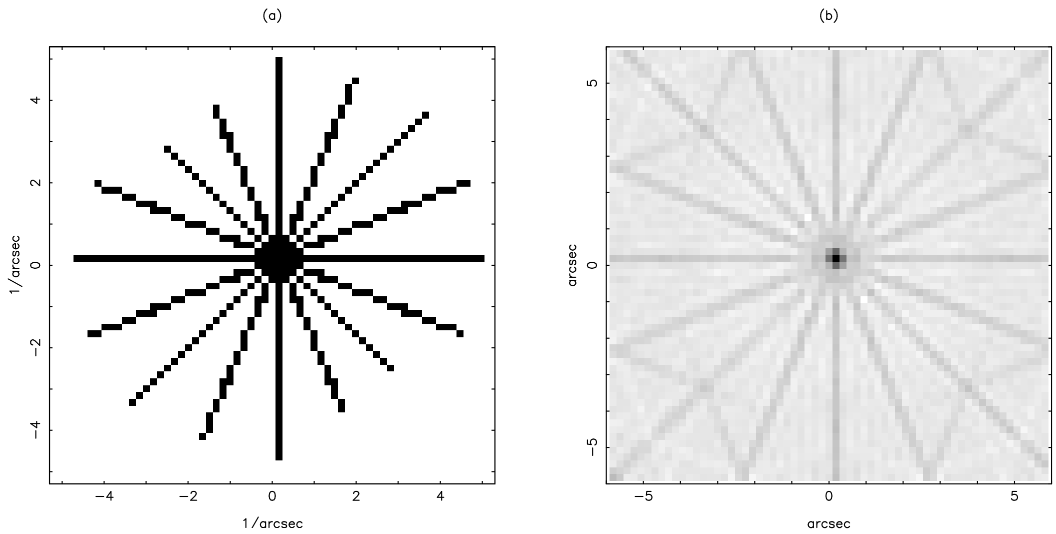

Figure 4.1: à droite, couverture fréquentielle de Hen 1379, et à gauche, la transformée de Fourier inverse des données (image originale).

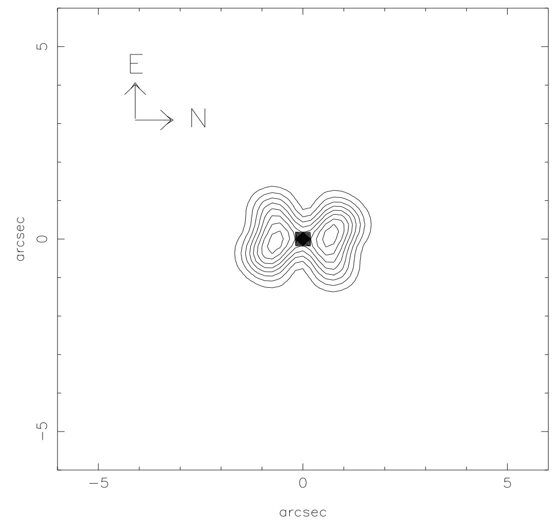

Figure 4.2: Hen 1379: image reconstruite.

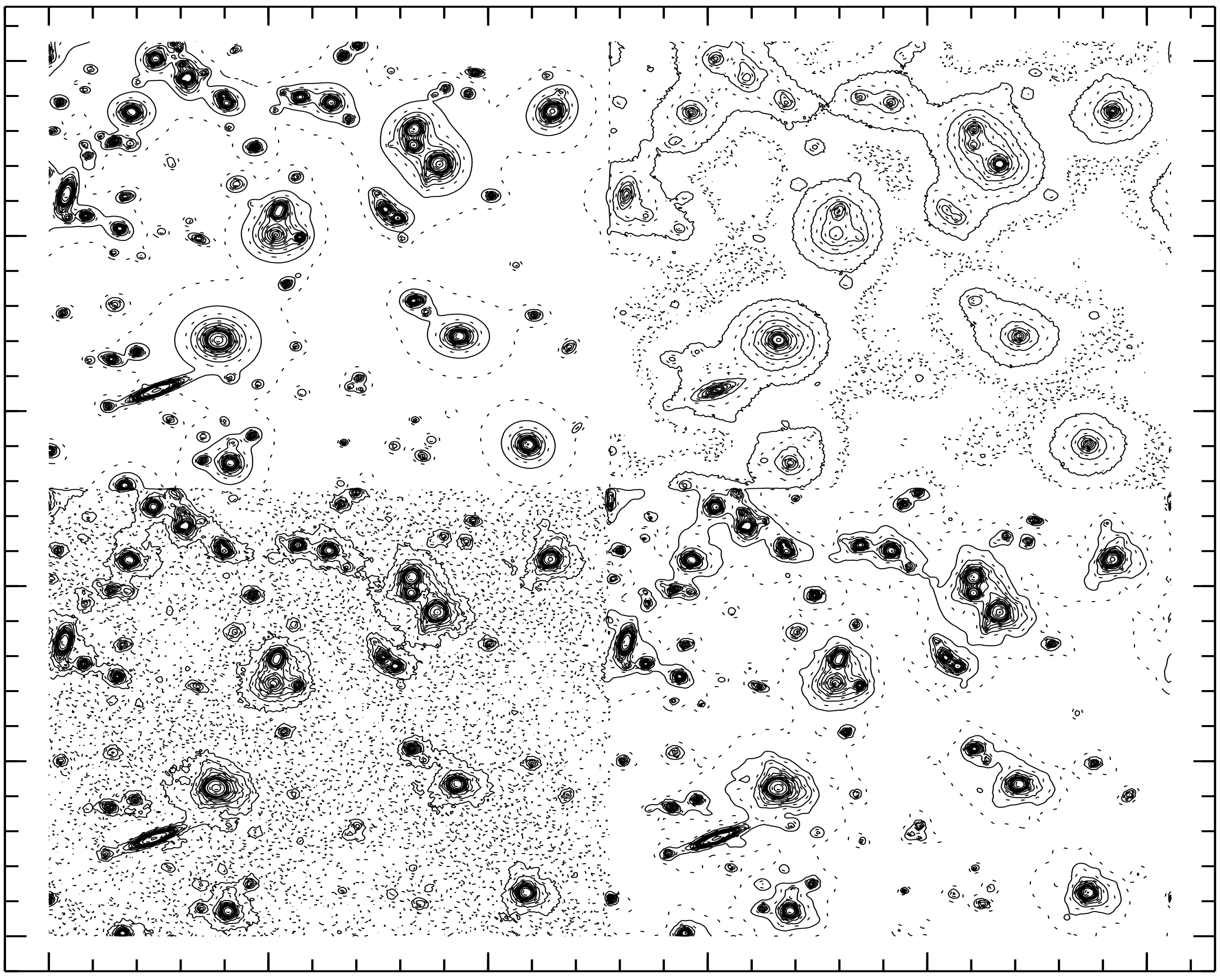

Figure 4.3: Image simulée de HST d'un amas distant de galaxies. En haut à gauche: image originale. En haut à droite: données convoluées et bruitées. En bas à gauche: restauration par la méthode de Richardson-Lucy non régularisée. En bas à droite: restauration par la méthode de Richardson-Lucy régularisée.

4.1.4 Compression d'images astronomiques

Avec l'arrivée d'un volume de données de plus en plus important (MEGACAM, VLT, ...) la question de la compression des données est devenue très importante. En collaboration avec le groupe chargé de l'archivage des données à l'ESO et le ST-ECF, nous avons développé une méthode de compression d'images astronomiques basée sur une approche pyramidale (Starck et al., 1996d; Starck et al., 1996e; Murtagh et al., 1998). Le principe de cette méthode est de sélectionner l'information que l'on veut garder en utilisant la transformée médiane pyramidale (PMT), et de ne coder que cette information, sans perte. La première étape cherche un nombre minimum de coefficients capable de produire une image de "haute qualité". La qualité est évidemment une notion subjective, et on définit par ce terme une image qui respecte les critères suivants:

- Il n'y a pas d'artefact dans l'image décompressée.
- Le résidu (image originale moins l'image décompressée) ne contient pas de structure.

La perte d'information est irréversible, et si on n'accepte pas la moindre perte, il faut alors aussi compresser le résidu. Dans ce cas, le taux de compression sera évidemment très faible (3 ou 4 au maximum). La méthode consiste à faire la séquence d'opérations suivante:

1. Détermination du support multirésolution.
2. Quantification des coefficients multirésolution produisant l'image filtrée.
3. Codage de chaque niveau de résolution en utilisant la méthode de Huang Bijaoui (1991). Elle consiste à appliquer un codage quadtree, suivi d'un codage de Huffman. Il n'y a pas d'information perdue au cours de cette phase.
4. Compression du bruit (si nécessaire).
5. La décompression consiste à reconstruire l'image filtrée et à rajouter la carte de bruit si elle a été compressée.

Comme la décompression se fait échelle par échelle (en partant de la résolution la plus faible) il n'est donc pas nécessaire de décompresser toute l'image pour faire une visualisation rapide.

Une étude a été menée au Centre de Données astronomiques de Strasbourg (CDS) dans le cadre du projet Aladin pour choisir une méthode de compression (Louys et al., 1999). L'étude tenait compte de la photométrie, de l'astrométrie des objets dans l'image décompressée, ainsi que de la détection d'objets. Une demi-douzaine de méthodes a été ainsi évaluée. La conclusion de cette étude est que la méthode basée sur la PMT permet d'obtenir des taux de compression impressionnant (jusqu'à 300 dans certains cas) avec une qualité équivalente à JPEG compressée à 40 (il avait auparavant été montré que JPEG pouvait permettre des taux de compression de 40, mais qu'au delà la dégradation photométrique devient inacceptable), tout en gardant une bonne qualité visuelle (ce qui n'est pas le cas de JPEG). Le choix du CDS s'est donc porté sur la PMT.

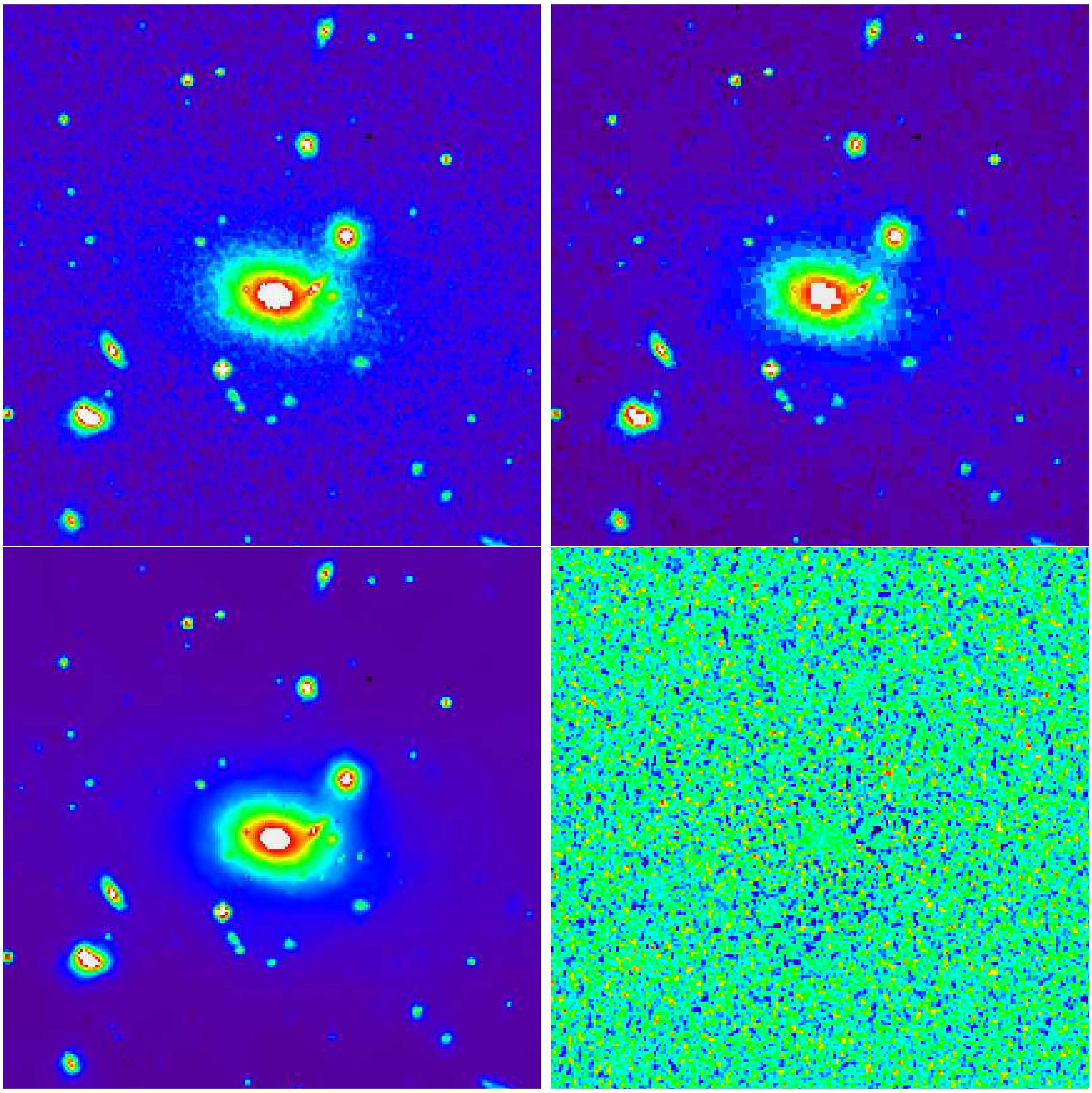

Figure 4.4: En haut à gauche, image HST de l'amas de galaxies de Coma. En haut à droite image décompressée en utilisant HCOMPRESS (30:1). En bas à gauche, image décompressée en utilisant PMT (30:1). En bas à droite, image originale moins l'image PMT décompressée.

4.1.5 Le projet ISO

Arrivé au CEA en 1994, peu avant le lancement du satellite infrarouge ISO (Infrared Space Observatory), mes travaux se sont naturellement orientés vers les problèmes de calibration et d'analyse des données ISO.

Le satellite ISO de l'Agence Spatiale Européenne a été lancé le 17 novembre 1995, avec à son bord quatre instruments dont la caméra ISOCAM. La caméra ISOCAM (Cesarsky et al., 1996) a été construite par un consortium de plusieurs laboratoires européens, le CEA étant maître d'oeuvre. Elle est composée de deux canaux, SW (2.2- 5.5 μm) et LW (4 - 18 μm), avec un détecteur de 32x32 pixels, une série de filtres, et quatre lentilles (1.5 à 12 secondes d'arc).

La calibration des données ISOCAM consiste en une série d'opérations:

- Normaliser les données en ADU (*analog-to-digital units*).
- Soustraire le *dark* (courant d'obscurité).
- Diviser par le *flat field* (champ uniforme).
- Co-additionner les données provenant de la même position du ciel.

En pratique, on est face à un certain nombre de problèmes:

- Le *dark* varie d'une orbite à une autre. La qualité du dark est donc médiocre et certains artefacts peuvent rester.
- la calibration des *flat field* est faite avec une source interne qui crée des *flats* différents de ceux de source astronomique.
- Les rayons cosmiques qui arrivent sur le détecteur et laissent des traces.
- Les effets de mémoire du détecteur. La transition entre deux niveaux d'intensité n'est pas instantanée.
- Les erreurs de pointage du satellite.
- La distorsion du champ de vue: la résolution de chaque pixel dépend de sa position dans le détecteur.
- La colonne 24 est morte.

Mon activité au sein du groupe ISOCAM s'est traduite par un certain nombre d'études (Starck et al., 1999a; et all., 1996; Altieri et al., 1999; Claret et al., 1999; Dubois et al., 1998; Starck et al., 1996b; Starck, 1996; Starck et al., 1996a; Claret and Starck, 1996; Djamdji et al., 1996; Starck et al., 1996c; Starck and Gastaud, 1995; Starck and Pantin, 1996; Siebenmorgen et al.,

1996) et la réalisation d'un certain nombre de programmes de réduction de données dans le logiciel CIA (*Cam Interactive Analysis*).

4.1.6 Multirésolution et données non astronomiques

En collaboration avec F. Murtagh, j'ai aussi exploré dans quelle mesure les algorithmes que nous avions développés pouvaient être directement applicables dans d'autres domaines. Ces travaux nous ont conduit à traiter sur des sujets aussi différents que le *clustering* (Murtagh and Starck, 1998; Murtagh and Starck, 1999), la prévision financière (Gonghuiand et al., 1999), la représentation spatiale de mesures économiques en agriculture (Morehart et al., 1999), les images vidéos (Daubos et al., 1999), ou les images médicales (Campbell et al., 1998).

4.2 Encadrement

J'ai co-encadré deux thésards au service d'astrophysique du CEA et encadré plusieurs stagiaires au cours des dernières années. J'ai aussi lancé (en 1996) et géré le projet *Multiresolution*, et animé la petite équipe CEA (deux personnes à temps complet, deux personnes à temps partiel, et les stagiaires) associée à ce projet.

4.2.1 Thèse de Eric Pantin

L'encadrement de la première thèse (15% de mon temps pendant deux ans), intitulée "Etude de la partie interne du disque de poussières de β pictoris: déconvolution des images à 10μm et modélisation de l'émission de poussières" et soutenue par E. Pantin en 1996, concernait la déconvolution. Les données infrarouge obtenues avec la caméra infrarouge TIMMI (placée sur le télescope de 3m60 de l'ESO) étaient particulièrement difficiles à dépouiller. Les effets conjugués de la PSF et d'un bruit assez important avaient rendu impossible l'analyse du disque sur les données brutes. Les méthodes traditionnelles de déconvolution ne permettaient pas de déconvoluer correctement à la fois le centre (très lumineux) et l'extérieur du disque. Nos travaux nous ont alors amenés à introduire le concept d'entropie multiéchelle (Pantin and Starck, 1996), et sont donc à l'origine de l'ensemble des travaux présentés aux second et troisième chapitres. L'utilisation de cette nouvelle forme d'entropie a ensuite permis de mettre en évidence une assymétrie dans le disque de l'étoile, et une des explications possibles pour comprendre cette assymétrie est la présence d'une planète (Lagage and Pantin, 1994; Pantin et al., 1997). L'image déconvoluée est présentée en section 3.2.5.

4.2.2 Thèse de Hervé Aussel

La seconde thèse, intitulée "Formation et évolution des galaxies, observées dans l'infrarouge moyen" et soutenue par Hervé Aussel en 1998, s'est déroulée dans le cadre du projet ISO.

L'utilisation de méthodes basées sur la multirésolution a permis l'exploitation automatique d'un grand nombre de données, et rendu possible la détection des objets les plus faibles. Les courbes $\log N - \log S$ (nombre de sources pour un flux donné) dérivées des données réduites par nos méthodes montrent que les modèles de formation de galaxies doivent faire intervenir une forte évolution à $z = 1$ (Aussel et al., 1999; Elbaz et al., 1999), ce qui n'est pas le cas avec les modèles standards.

La principale difficulté d'ISOCAM concernant la détection de sources faibles est la combinaison de rayons cosmiques et d'effets mémoires du détecteur (Starck et al., 1999a). En effet, quand un rayon cosmique produit une simple montée puis descente du signal, il peut être éliminé relativement bien par un filtrage médian. Mais quand ce n'est pas le cas, les effets mémoires peuvent produire de fausses détections. La principale source d'erreur n'est donc pas la limite de détection de l'instrument mais le grand nombre de rayons cosmiques qui créent de fausses détections.

Trois types de rayons cosmiques peuvent être distingués:

- une structure positive, forte, et courte en durée.
- une structure positive, suivie d'une queue positive.
- une structure positive, suivie d'une queue négative.

La figure 4.5 montre le signal mesuré (en ADU) par un pixel en fonction du temps (fréquence d'échantillonnage = 2.1s). On peut y voir les trois types de rayon cosmique.

Pour pouvoir détecter les objets les plus faibles, il est nécessaire de pouvoir corriger le signal de ces effets de mémoire consécutifs aux rayons cosmiques. Pour cela, nous avons mis au point une méthode de calibration à partir de la reconnaissance de formes (Starck et al., 1999b; Aussel et al., 1999). Elle consiste à décomposer un signal suivant les objets qui le composent par un modèle de vision multi-échelle (Bijaoui and Rué, 1995; Starck et al., 1997), puis à rejeter les objets ne vérifiant pas certaines règles.

La figure 4.6 présente le résultat après un tel traitement. Les données originales sont montrées en haut. La figure 4.7 montre la décomposition du signal original (voir Figure 4.6 en haut) en ses principaux constituants: (a), (b), et (d) sont les structures (rayons cosmiques courts, rayon cosmique avec queue négative, et ligne de base) qui ne présentent pas d'intérêt pour la détection de sources, et (c) et (e) (source et bruit) doivent être utilisés plus loin dans la chaîne de traitement. Le bruit doit aussi être gardé car des sources faibles indétectables dans le signal temporel peuvent le devenir une fois toutes les données co-additionnées. La simple somme de ces cinq composants reproduit le signal original. Les données calibrées et sans fond de ciel sont obtenues par addition des signaux (c) et (e).

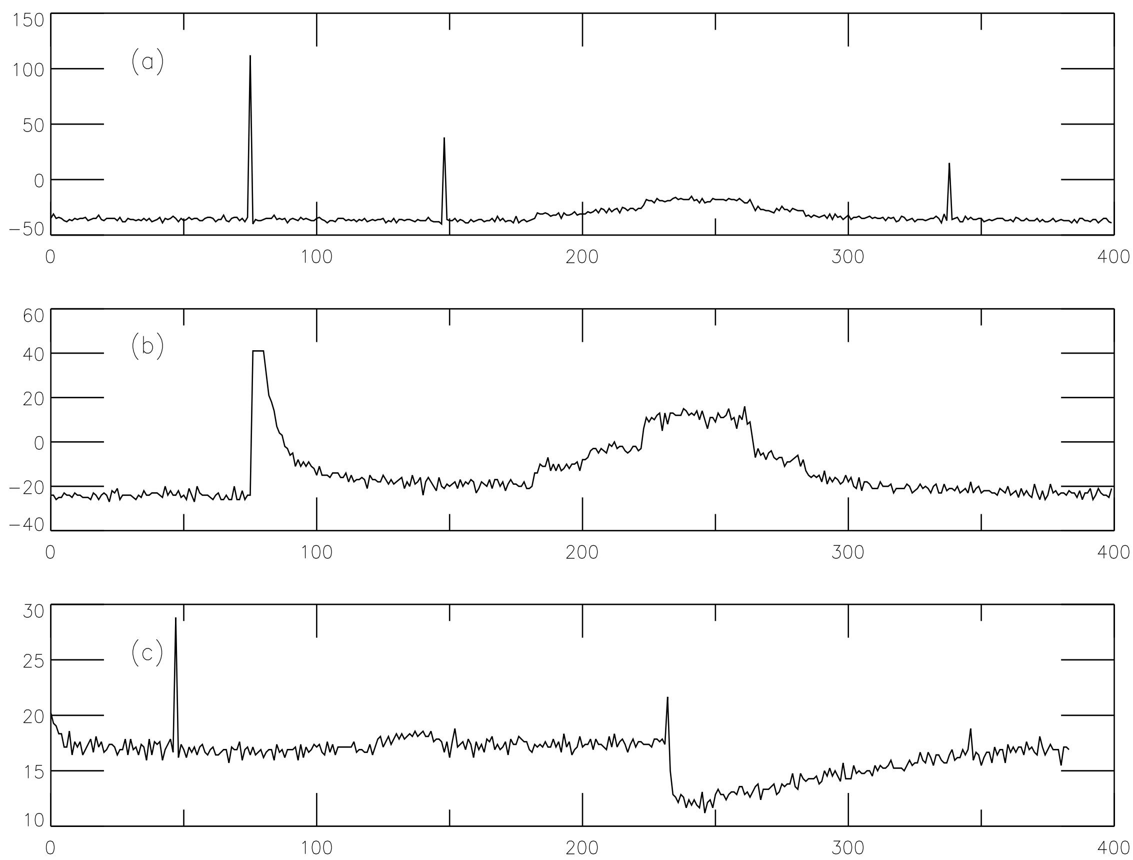

Figure 4.5: Exemples de rayons cosmiques. (a) Ce signal contient trois rayons cosmiques du premier type. (b) Un rayon cosmique avec une longue queue apparaît (appelé *fader*) autour de la position 80. Le rayon cosmique a été tronqué, et sa réelle amplitude est 2700 ADUs. (c) Un rayon cosmique avec une queue négative (appelé *deeper*) apparaît autour de la position 240.

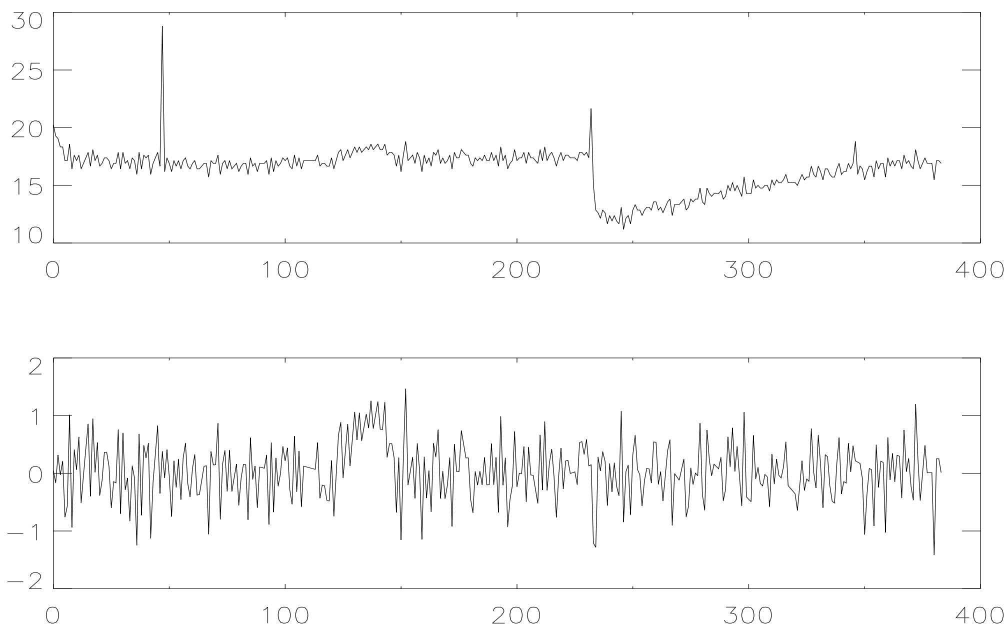

Figure 4.6: En haut, données originales, et en bas, données calibrées. On peut remarquer la variation de gain d'environ 5 ADUs qui apparaît après le second rayon cosmique.

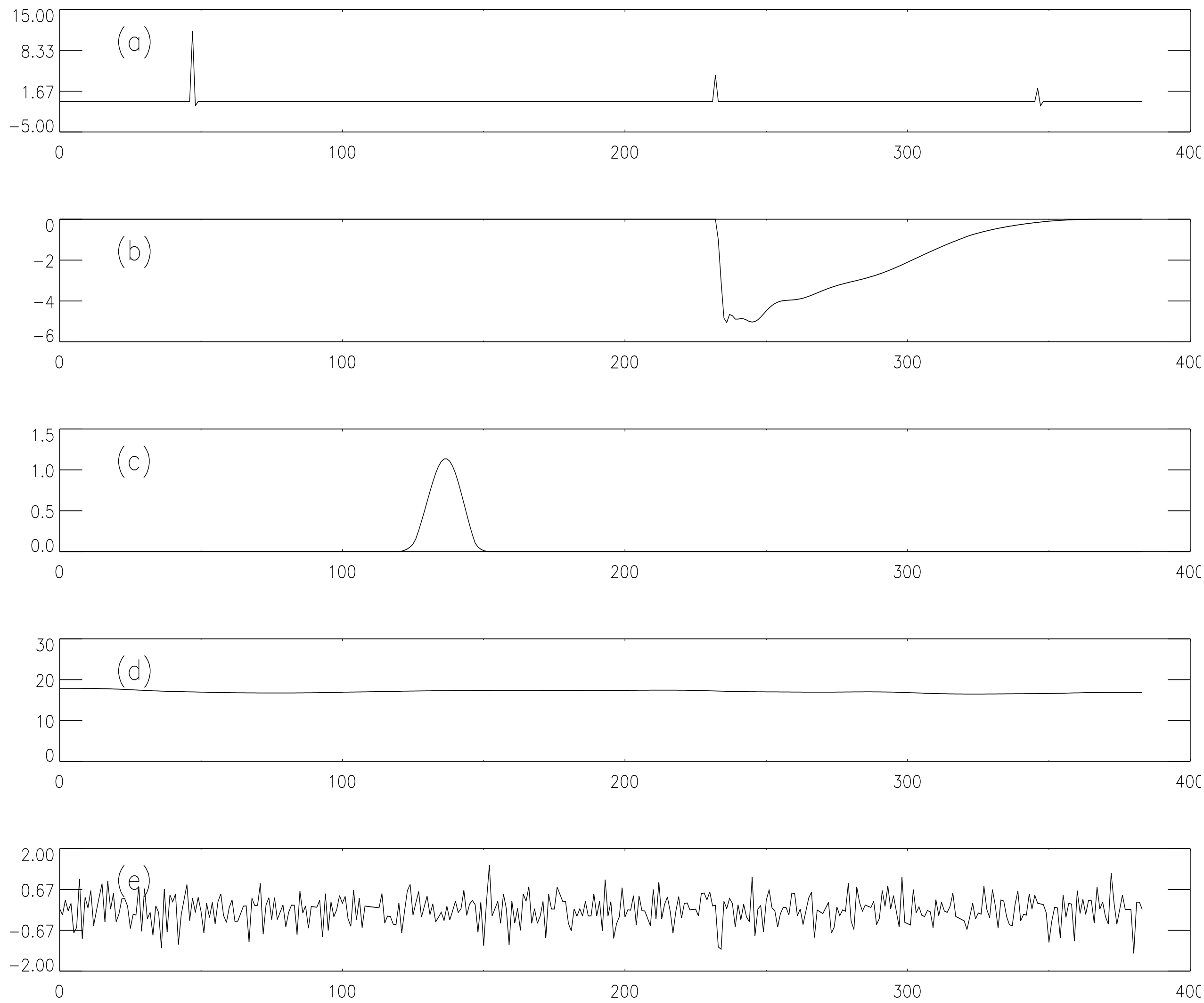

Figure 4.7: Décomposition d'un signal suivant ses principaux constituants: (a) rayons cosmiques courts, (b) le rayon cosmique avec la queue négative, (c) source, (d) ligne de base, (e) bruit. La simple somme de ces cinq composants reproduit le signal original. Les données calibrées et sans fond de ciel sont obtenues par addition des signaux (c) et (e).

4.2.3 Stages concernant l'analyse de la structure de l'émission X

Afin de préparer la mission XMM, une série de stages (C. Delattre, J. Marichez, M. Albertelli, T. Douvion) a été effectuée pour développer et tester, à l'aide de simulations, des algorithmes de détection d'objets dans des images à très faible flux de photons. Ces algorithmes ont été appliqués sur des simulations, et dans le cadre d'une collaboration avec M. Pierre du SAP, sur des données X ROSAT (HRI) afin de mettre en évidence des structures au sein des amas des galaxies. Les résultats sont présentés et comparés avec ceux obtenus par des méthodes plus traditionnelles (voir les figures 4.8,4.9, 4.10). Nous avons publié la méthode (Starck and Pierre, 1998) et l'avons appliqué à une série de 23 amas situés entre $0.01 < z < 0.32$ (Pierre and Starck, 1998).

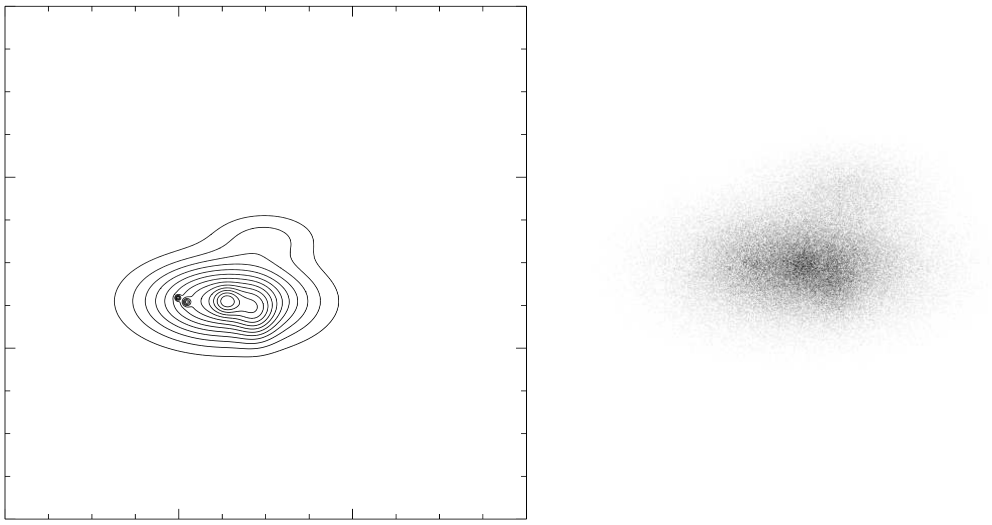

Figure 4.8: (gauche) Simulation d'un amas constitué d'un corps principal elliptique et comprenant les structures suivantes: un courant de refroidissement au centre, une faible sous-structure à droite du centre, deux sources ponctuelles à gauche et un groupe au-dessus en train de fusionner avec l'amas. (droite) l'image avec un bruit poissonnien.

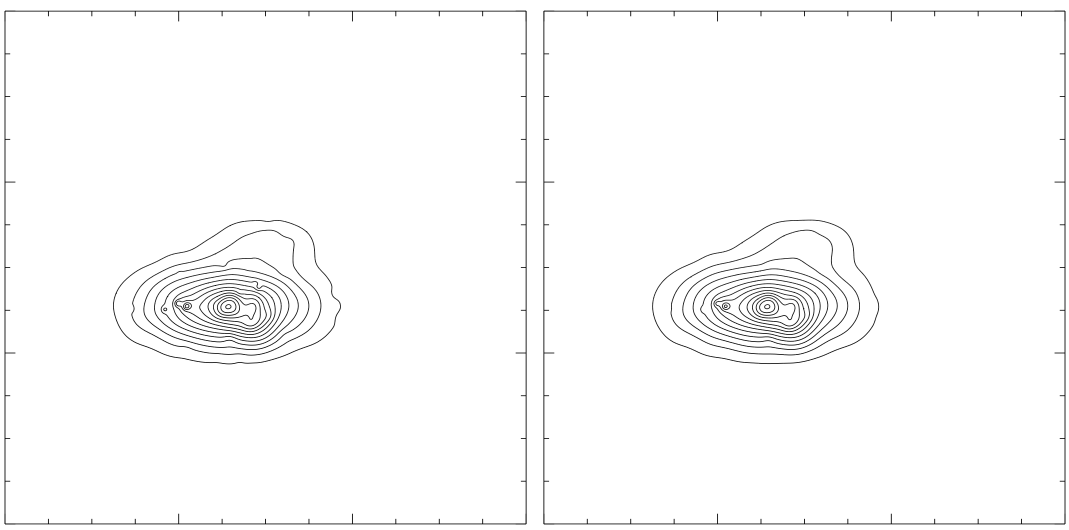

Figure 4.9: Filtrage de l'image 4.8 selon la méthode ondelette + bruit poissonnien; (gauche) niveau de confiance 10^{-3}, (droite) 10^{-4}.

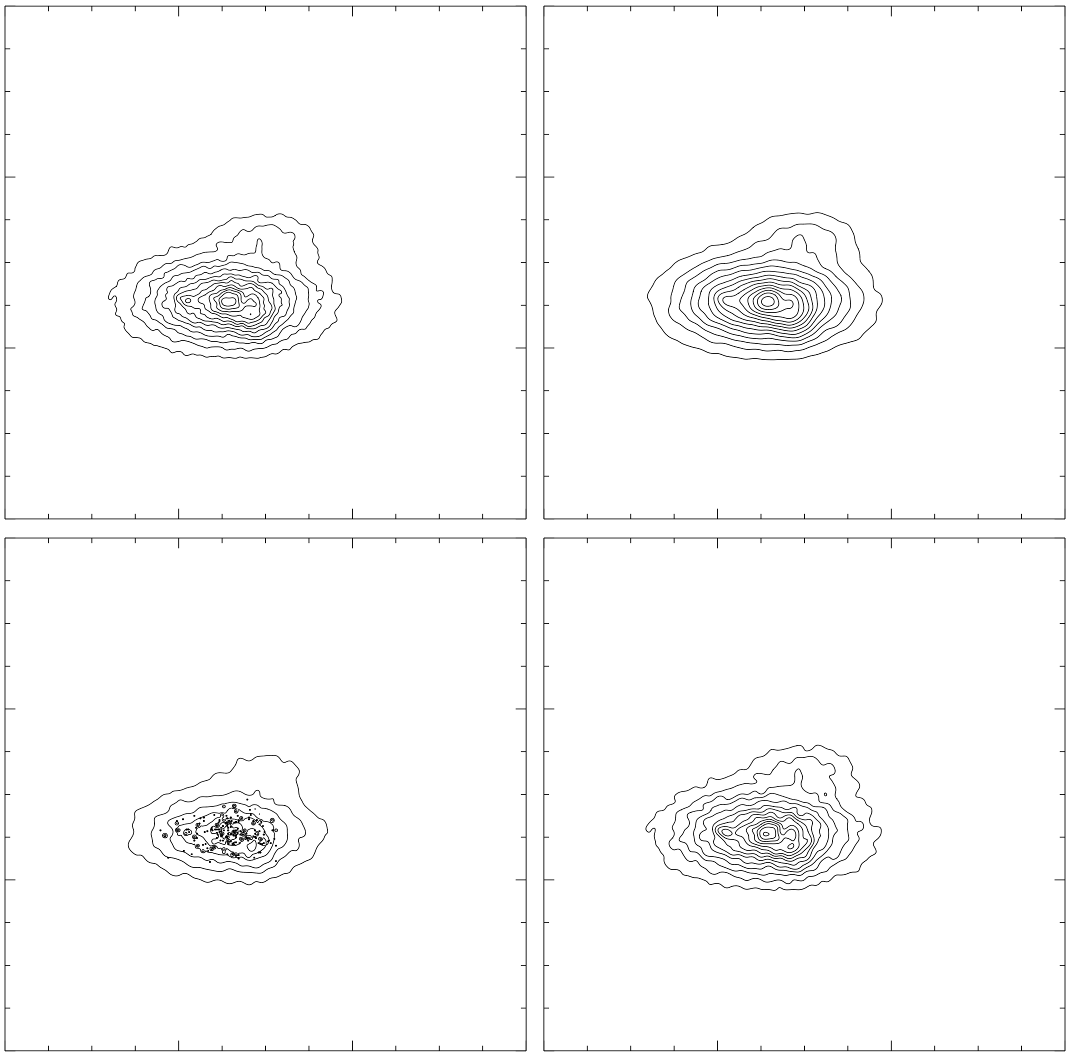

Figure 4.10: (haut) L'image 4.8 traitée avec un filtre Gaussien traditionnel ($\sigma = 4$ et 7 pixels). (en bas à gauche) filtre ondelette avec une rejection à 3 sigma sur chaque échelle ondelette puis détection à 10σ; (en bas à droite) ondelettes + hypothèse de bruit Gaussien et une détection à 10σ.

Analyse de l'amas de galaxies A2390

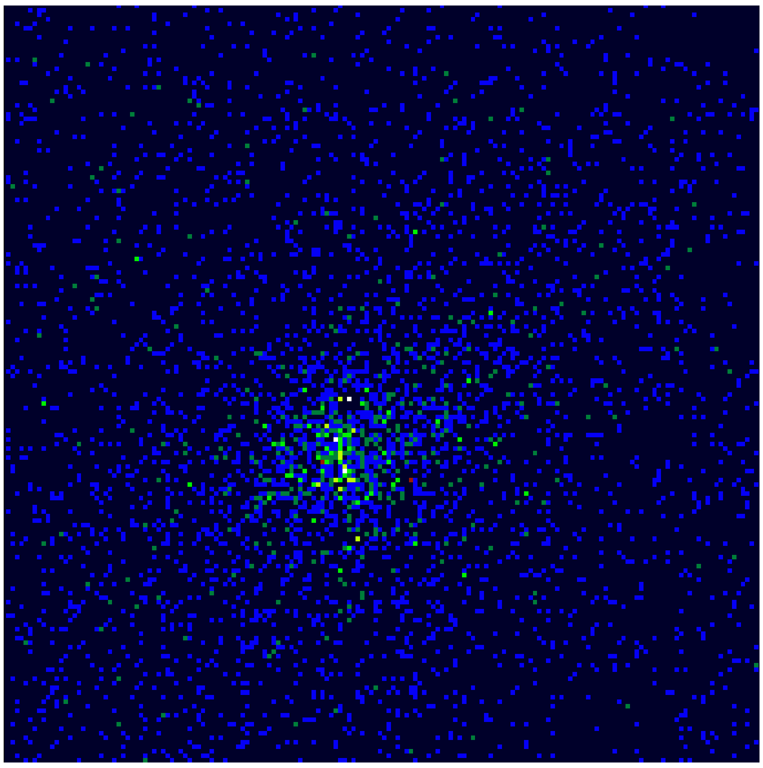

Figure 4.11: Image ROSAT de l'amas A2390.

L'amas de galaxies A2390 est situé à un redshift de 0.231. La figure 4.11 montre une image de cet amas, obtenue avec le satellite X ROSAT. La résolution est d'une seconde d'arc par pixel, avec un nombre total de photons égal à 13506 pour un temps d'exposition de huit heures et demi. Le niveau de fond est d'environ de 0.04 photons par pixel. Il est évident que cette image brute est peu exploitable, et il est nécessaire de la traiter pour en extraire de l'information. La méthode standard consiste à convoluer l'image par une gaussienne. La figure 4.12 montre le résultat de cette opération (Gaussienne de largeur à mi-hauteur de 5", qui est à peu près celle de la réponse instrumentale de l'instrument). L'image lissée montre des structures, mais on voit aussi qu'il reste beaucoup de bruit, et il est difficile de donner une signification à ces structures. La figure 4.13 montre l'image filtrée par une méthode basée sur les ondelettes. Le bruit a été éliminé, et on voit que l'analyse en ondelettes a permis de mettre en évidence des structures très faibles dans l'émission X, permettant d'expliquer des phénomènes d'amplification gravitationnelle observés dans le domaine visible.

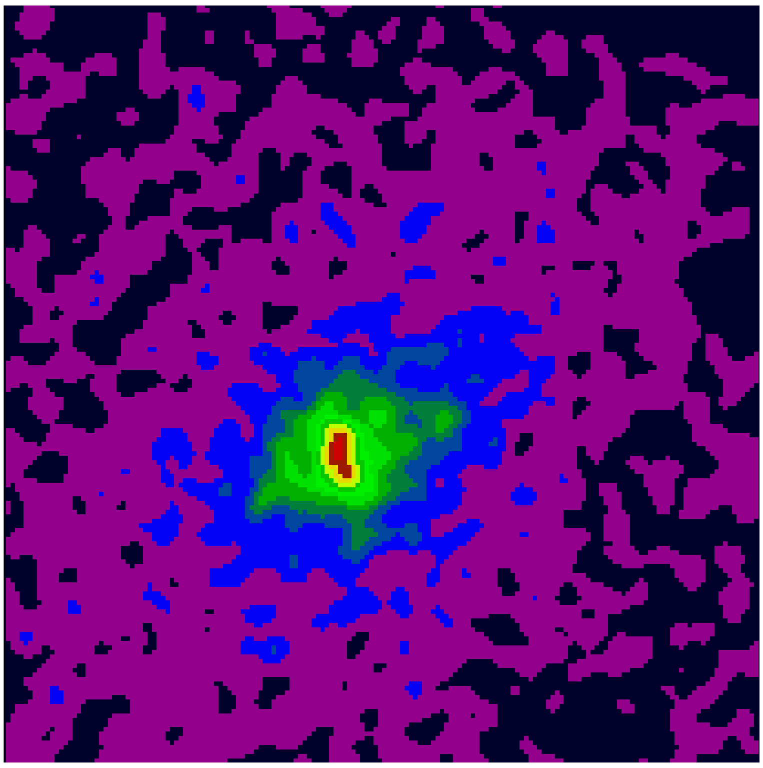

Figure 4.12: Image ROSAT de l'amas A2390 filtrée par la méthode standard (convolution par une Gaussienne).

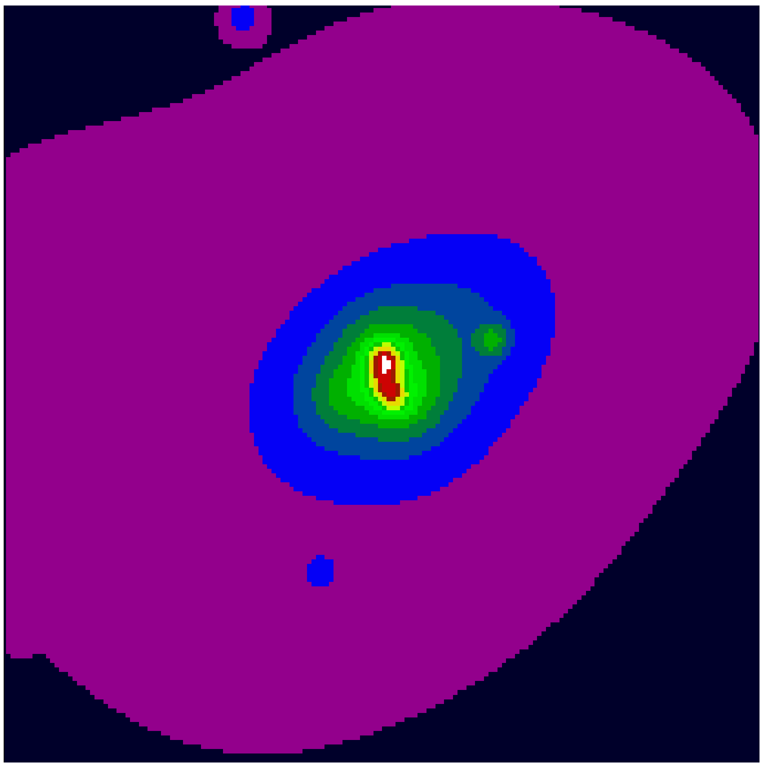

Figure 4.13: Image ROSAT de l'amas A2390 filtrée par la méthode basée sur les ondelettes.

4.2.4 Stages concernant la déconvolution

J'ai également encadré deux stages (N. Fisher et R. Bourgon) sur des développements sur l'entropie multiéchelle, et un troisième (O. Dubois) sur des aspects de l'utilisation d'un modèle markovien pour la déconvolution des images astronomiques (collaboration avec J. Idier du L2S).

4.2.5 Stages concernant une estimation du bruit dans les images

Une évaluation des différentes méthodes d'estimation de l'écart type d'un bruit Gaussien a été réalisée par A. Lebras, et les implications sur la détection d'objet de l'erreur sur cette estimation a été étudiée. Cette étude a permis en particulier de d'établir qu'une erreur de moins de 5% sur l'estimation de l'écart type du bruit avait une répercution négligeable sur la détection des objets par les méthodes multiéchelles.

4.2.6 Le projet Multirésolution

Le but de ce projet était de développer un système de traitement de données par analyse en multirésolution. Pour le réaliser, il fallait reprendre le travail effectué lors de mon passage à l'ESO (contexte ondelette), l'enrichir des travaux les plus récents dans le domaine. La vocation d'un tel système est double, tout d'abord de disposer d'outils pointus pour le traitement des données, puis de servir de plate-forme de développement et d'intégration de nouvelles méthodes spécifiques aux besoins des projets du service d'astrophysique. Pour réaliser ces objectifs, une collaboration s'est faite entre le CEA et l'Observatoire de la Côte d'Azur, une équipe a été formée, et la première version du logiciel MR/1 a été finie en juin 1997. Elle est disponible sur PC (linux et windows), SUN, DEC-OSF, IBM, et HP. Les codes sont en C++ et JAVA (environ 80000 lignes de code+200 pages de documentation).

Les différents programmes composant MR/1 sont:

- Les transformées en ondelettes: de nombreux algorithmes ont été développés, permettant d'adapter l'analyse de manière optimale.

- La compression d'images: une méthode s'appuyant sur la multirésolution a été créée, celle-ci permet d'avoir un taux de compression important en gardant une bonne qualité visuelle, astrométrique, et photométrique sur les objets contenus dans l'image. Cette méthode permet aussi de ne décompresser l'image qu'a des résolutions inférieures, ce qui est un gros avantage pour la transmission par réseau d'images contenues dans une base de données. En effet, l'utilisateur peut très rapidement voir une image à très mauvaise résolution, et si l'image l'intéresse, continuer à rapatrier les résolutions suivantes (ce que l'on appelle transmission progressive). Une étude a été menée au Centre de Données de Strasbourg (CDS) montrant qu'un taux de compression de 300 pouvait être atteint avec une qualité visuelle d'image acceptable pour les images obtenues avec des plaques de Schmidt de l'ESO, digitalisées par la MAMA (Machine Automatique à Mesurer pour l'Astronomie). La norme JPEG ne permettait qu'un taux de compression de 40 pour la même qualité visuelle.

- La déconvolution d'images: la déconvolution consiste à tenter de retirer l'effet de l'instrument (télescope) sur les données. Cet effet se traduit par une perte de résolution. Le bruit présent dans les données rend le problème insoluble (pas de solution unique) si on n'ajoute pas de contraintes. Nos travaux ont montré que ces contraintes pouvaient s'exprimer dans l'espace des ondelettes, et donc de mieux tenir compte de la statistique du bruit.

- L'analyse spectrale par ondelettes. Les ondelettes ont apporté une nouvelle approche pour l'analyse des données. La détection de bande d'absorption à faible rapport signal sur bruit a pu être possible, de même que l'extraction de paramètres tel que l'épaisseur optique du milieu.

- L'analyse de données X (ROSAT) par ondelettes. Les sources peuvent être détectées de manière automatique, et les images restaurées permettent aussi de faire des analyses morphologiques d'amas de galaxies.

- La détection automatique de sources dans les données. Un modèle de vision multi-échelle permet de détecter et d'extraire les objets ponctuels et étendus compris dans une image.

- Le recalage d'images: dans le cas où l'on dispose de deux images, on peut les recaler de manière automatique, en se basant sur les structures détectées dans l'espace ondelette.

MR/1 a été très largement utilisé dans le cadre du projet ISO, à travers ses programmes de déconvolution et de détection automatique d'objets, et aussi par l'utilisation de classes C++ pour bâtir des programmes dédiés à ISO. La double vocation du projet a donc bien été une réussite. La challenge de MR/1 est maintenant de pouvoir servir les futurs projets du SAP, tels que XMM, MEGACAM, ou PLANCK.

4.3 Enseignement

J'ai enseigné en 1998 et en 1999 en dernière année d'école d'ingénieur (TELECOM-INT). La première partie du cours traite des différentes formes de transformée multiéchelle, transformée en ondelettes continue, transformée bi-orthogonale, algorithme à trous, transformées basées sur la morphologie mathématique, etc. Ensuite la distribution de probabilité d'un coefficient d'ondelettes est étudiée pour plusieurs modèles de bruit (bruit blanc Gaussien, Poissonnien, speckle, bruit corrélé, etc), et le concept de support multirésolution est introduit. La dernière partie concerne les applications. On voit comment une transformée multiéchelle associée à un modèle de bruit peut être utilisée pour des applications aussi diverses que

- le filtrage.
- la déconvolution.
- la compression d'image.
- le recalage d'image
- la détection et l'extraction d'objet.
- la détection de contour.
- l'analyse multifractale.

De plus, j'ai donné un certain nombre de séminaires, un *tutorial* (conférence ADASS) sur l'analyse en multirésolution en astronomie, et j'ai participé à deux écoles d'été, la première aux Houches (*on Infrared Space Astronomy, Today and Tomorrow*) où j'ai présenté en détail comment traiter les données ISOCAM (calibration des données, détection d'objet, déconvolution), et la seconde à Pune en Inde (*Data Visualization and Image Processing Workshop*) où j'ai fait un cours sur les ondelettes.

Bibliographie

Aczel and Daroczy, 1978, On Measures of Information and their Characterizations, Academic Press

Adelson, E., Simoncelli, E., and Hingorani, R., 1987, SPIE Visual Communication and Image Processing II 845, 50

Altieri, B., Metcalfe, L., Blommaert, J., Elbaz, D., Starck, J.-L., and Aussel, H., 1999, Experimental Astronomy, submitted

Amato, U. and Vuza, D., 1997, Numerical Functional Analysis and Optimization 18, 461

Antonini, M., Barlaud, M., Mathieu, P., and Daubechies, I., 1992, IEEE Transactions on Image Processing 1(2), 205

Aussel, H., Elbaz, D., Cesarsky, C., and Starck, J., 1999, Astronomy and Astrophysics 342, 313

Bijaoui, A., 1984, Introduction au Traitement Numérique des Images, Masson, Paris

Bijaoui, A. and Rué, F., 1995, Signal Processing 46, 229

Bijaoui, A., Starck, J., and Murtagh, F., 1994, Traitement du Signal 3, 11

Blanc-Féraud, L. and Barlaud, M., 1996, Vistas in Astronomy 40(4), 531

Bontekoe, T., Koper, E., and Kester, D., 1994, Astronomy and Astrophysics. 294, 1037

Burg, J., 1978, Annual Meeting International Society Exploratory Geophysics, Reprinted in Modern Spectral Analysis, D.G. Childers, ed., IEEE Press, New York, 34–41

Campbell, J.-G., Murtagh, F., and Starck, J.-L., 1998, in Bioengineering in Ireland & Ulster Biomedical Engineering Society Conference, Dublin

Cesarsky, C., Abergel, A., Agnese, P., et al., 1996, Astronomy and Astrophysics 315, L32

Chambolle, A., DeVore, R., Lee, N., and Lucier, B., 1998, IEEE Transactions on Signal Processing 7(3), 319

Champagnat, F., Goussard, Y., and Idier, J., 1996, IEEE Transactions on Image Processing 44(12), 2988

Charter, M., 1989, in Maximum Entropy and Bayesian Methods, pp 325–339, Kluwer

Chipman, H., Kolaczyk, E., and McCulloch, R., 1997, Journal of the American Statistical Association 92(440), 1413

Cittert, P. V., 1931, Zeitschrift für Physik 69, 298

Claret, A., Dzitko, H., Engelmann, J., and Starck, J.-L., 1999, Experimental Astronomy, in press

Claret, A. and Starck, J., 1996, ISOCAM Data Analysis with X_CIA Environment, Technical report, CEA

Cohen, A., Daubechies, I., and Feauveau, J., 1992, Communications in Pure and Applied Mathematics 45, 485

Crouse, M., Nowak, R., and Baraniuk, R., 1998, IEEE Transactions on Signal Processing 46, 886

Csillaghy, A., 1996, Vistas in Astronomy 40, 503

Daubechies, I., 1992, Ten Lectures on Wavelets, Society for Industrial and Applied Mathematics (SIAM), Philadelphia

Daubos, T., Geradts, Z., Starck, J., Campbell, J., and Murtagh, F., 1999, in SPIE conference on Signal and Image Processing: Wavelet Applications in Signal and Image Processing VII, Denvers, 18-23 July

Demoment, G., 1989, IEEE Transactions on Acoustics Speach and Signal Processing 37(12), 2024

Djamdji, J., Starck, J., and Claret, A., 1996, ISOCAM Images Registration, Technical report, CEA

Donoho, D., 1993, Proceedings of Symposia in Applied Mathematics 47

Donoho, D. and Johnstone, I., 1993, Ideal spatial adaptation by wavelet shrinkage, Technical Report 400, Stanford University

Dubois, O., Starck, J., and Idier, J., 1998, Deconvolution of astronomical images using a Markov random, Technical report, CEA

Elbaz, D., Cesarsky, C., Fadda, D., Aussel, H., Désert, F., Franceschini, A., Flores, H., Harwit, M., Puget, J., Starck, J., Danese, L., Koo, D., and Mandolesi, R., 1999, Astronomy and Astrophysics, submitted

et all., R. S., 1996, Astronomy and Astrophysics 315, L169

Feauveau, J., 1990, Ph.D. thesis, Université Paris Sud

Frieden, B., 1978a, Image Enhancement and Restoration, Springer-Verlag, Berlin

Frieden, B., 1978b, Probability, Statistical Optics, and Data Testing: a problem solving approach, John Wiley & Sons, Berlin, springer-verlag edition

Galatsanos, N. and Katsaggelos, A., 1992, IEEE Transactions on Image Processing 1, 322

Gastaud, R., Starck, J.-L., and Pierre, M., 1999, in D. M. Mehringer, R. L. Plante, and D. A. Roberts (eds.), Astronomical Data Analysis Software and Systems VIII, Vol. 172, pp 399–402

Geman, D. and Reynolds, G., 1992, IEEE Transactions on Pattern Analysis and Machine Intelligence 14(3), 367

Golub, G., Heath, M., , and Wahba, G., 1979, Technometrics 21, 215

Gonghuiand, Z., Starck, J., Campbell, J., and Murtagh, F., 1999, Journal of Computational Intelligence in Finance 7(3)

Gull, S. and Skilling, J., 1991, MEMSYS5 Quantified Maximum Entropy User's Manual

Högbom, J., 1974, Astronomy and Astrophysics Supplement Series 15, 417

Holschneider, M., Kronland-Martinet, R., Morlet, J., and Tchamitchian, P., 1989, Wavelets: Time-Frequency Methods and Phase-Space, Chapt. A Real-Time Algorithm for Signal Analysis with the Help of the Wavelet Transform, pp 286–297, Springer-Verlag

Huang, L. and Bijaoui, A., 1991, Experimental Astronomy 1, 311

Hunt, B., 1994, Int. Journal of Modern Physics C 5(1), 151

Jansson, P., Hunt, R., and Peyler, E., 1970, Journal of the Optical Society of America 60, 596

Jaynes, E., 1957, Phys. Rev. 106, 171

Johnson, R. and Shore, J., 1984, Proc. IEEE Int. Conf. Acoust. Speech., Signal Processing 32, 129

Kaaresen, K., 1997, IEEE Transactions on Image Processing 45(5), 1173

Krim, H., Drouiche, K., and Pesquet, J. C., 1992, in Time-Frequency and Time-Scale Analysis, pp 105–108, IEEE Signal Processing Society, Victoria, B.C., Canada

Lagage, P. and Pantin, E., 1994, Nature 369, 628+

Landweber, L., 1951, American Journal of Mathematics 73, 615

Lannes, A. and Roques, S., 1987, Journal of the Optical Society of America 4(1)

Louys, M., Starck, J.-L., Mei, S., Bonnarel, F., and Murtagh, F., 1999, Astronomy and Astrophysics, Suppl. Ser. 136, 579

Lucy, L. B., 1974, Astronomical Journal 79, 745

Magain, P., Courbin, F., and Sohy, S., 1998, Astrophysical Journal 494, 472

Mallat, S., 1998, A Wavelet Tour of Signal Processing, Academic Press

Mallat, S. G., 1989, IEEE Transactions on Pattern Analysis and Machine Intelligence 11(7), 674

Mohammad-Djafari, A., 1994, Traitement du Signal pp 87–116

Mohammad-Djafari, A., 1998, Traitement du Signal 15(6), 545

Morehart, M., Murtagh, F., and Starck, J., 1999, International Journal of Geographical Information Science 13(6), 557

Moulin, P. and Liu, J., 1999, IEEE Transactions on Information Theory 45, 909

Murtagh, F. and Starck, J., 1998, Pattern Recognition 31(7), 847

Murtagh, F. and Starck, J., 1999, Computer Journal, submitted

Murtagh, F., Starck, J., and Bijaoui, A., 1995a, ST-ECF Newsletter 22, 6

Murtagh, F., Starck, J., and Bijaoui, A., 1995b, Astronomy and Astrophysics, Suppl. Ser. 112, 179

Murtagh, F., Starck, J., and Louys, M., 1998, International Journal of Imaging Systems and

Technology 9, 38

Murtagh, F., Zeilinger, W., and Starck, J., 1993, ESO Messenger 73, 37

Narayan, R. and Nityananda, R., 1986, Ann. Rev. Astron. Astrophys. 24, 127

Nason, G., 1996, Journal of the Royal Statistical Society B 58, 463

Nason, G. P., Mar. 24, 1994, Wavelet regression by cross-validation, Technical report, Dept. of Math., Univ. of Bristol, University Walk, Bristol, BS8 1TW, U.K.

Nunez, J. and Llacer, J., 1998, AAS 131, 167

Olsen, S., 1993, CVGIP 55, 319

P. Cruzalèbes, G. S. and Starck, J., 1992, Journal of the Optical Society of America A 9(5)

Pantin, E., Lagage, P. O., and Artymowicz, P., 1997, Astronomy and Astrophysics 327, 1123

Pantin, E. and Starck, J., 1996, Astronomy and Astrophysics, Suppl. Ser. 315, 575

Pierre, M. and Starck, J., 1998, Astronomy and Astrophysics 128(3), 801

Puetter, R., 1986, Int. J. Image Sys & Tech. 6, 314

Richardson, W., 1972, Journal of the Optical Society of America 62, 55

Said, A. and Pearlman, W., 1996, IEEE Trans. on Circ. and Syst. for Video Tech. 6(3), 243

Shannon, C., 1948, Bell System Tech J. 27, 379

Shapiro, J., 1993, IEEE Trans. on Signal Process. 41(12), 3445

Shore, J. and Johnson, R., 1981, IEEE Transactions on Information Theory 26, 472

Siebenmorgen, R., Starck, J., Cesarsky, D., Guest, S., and Sauvage, M., 1996, ISOCAM Data Users Manual Version 2.1, Technical Report SAI/95-222/Dc, ESA

Skilling, J., 1989, in Maximum Entropy and Bayesian Methods, pp 45–52, Kluwer

Starck, J., 1993, The Wavelet Context, Technical report, European Southern Observatory – Image Processing Group

Starck, J., 1996, Analysis of VLA Lockman Hole data, Technical report, CEA

Starck, J., Abergel, A., Aussel, H., Sauvage, M., Gastaud, R., Claret, A., Desert, X., Delattre, C., and Pantin, E., 1999a, Astronomy and Astrophysics, Suppl. Ser. 134, 135

Starck, J., Aussel, H., and Elbaz, D., 1996a, Sources Detection with ISOCAM, Technical report, CEA

Starck, J., Aussel, H., Elbaz, D., Fadda, D., and Cesarsky, C., 1999b, Astronomy and Astrophysics, Suppl. Ser. 138, 365

Starck, J., Aussel, H., Gastaud, R., and Elbaz, D., 1996b, Deep Survey Data Calibration from Pattern Recognition, Technical report, CEA

Starck, J. and Bijaoui, A., 1994a, Signal Processing 35, 195

Starck, J. and Bijaoui, A., 1994b, Journal of the Optical Society of America A 11(4)

Starck, J., Bijaoui, A., Lopez, B., and Perrier, C., 1994, Astronomy and Astrophysics 283, 349

Starck, J., Bijaoui, A., and Murtagh, F., 1995, CVGIP: Graphical Models and Image Processing 57, 420

Starck, J., Claret, A., and Siebenmorgen, R., 1996c, ISOCAM Data Calibration, Technical report, CEA

Starck, J. and Gastaud, R., 1995, ISOCAM Interactive Analysis: an overview, Technical report, CEA

Starck, J. and Murtagh, F., 1994, Astronomy and Astrophysics 288, 343

Starck, J. and Murtagh, F., 1998, Publications of the Astronomical Society of the Pacific 110(744), 193

Starck, J. and Murtagh, F., 1999, Signal Processing 76(2), 147

Starck, J., Murtagh, F., and Bijaoui, A., 1998a, Image Processing and Data Analysis: The Multiscale Approach, Cambridge University Press, Cambridge (GB)

Starck, J., Murtagh, F., Bonnarel, F., and Couvidat, S., 1999c, IEEE Transaction on Information Theory, submitted

Starck, J., Murtagh, F., and Durand, D., 1996d, NASA Science Information Systems Newsletter 2(39)

Starck, J., Murtagh, F., and Gastaud, R., 1998b, Special Issue on Multirate Systems, Filter Banks, Wavelets, and Applications of IEEE Transactions on CAS II 45(8)

Starck, J., Murtagh, F., Pirenne, B., and Albrecht, M., 1996e, Publications of the Astronomical Society of the Pacific 108, 446

Starck, J. and Pantin, E., 1996, Second Order Dark Correction, Technical report, CEA

Starck, J. and Pierre, M., 1998, Astronomy and Astrophysics, Suppl. Ser. 128

Starck, J., Siebenmorgen, R., and Gredel, R., 1997, Astrophysical Journal 482, 1011

Tikhonov, A., Goncharski, A., Stepanov, V., and Kochikov, I., 1987, Soviet Physics – Doklady 32, 456

Timmermann, K. E. and Nowak, R., 1999, IEEE Transactions on Signal Processing 46, 886

Titterington, D., 1985, AA 144, 381

Vidakovic, B., 1998, Journal of the American Statistical Association 93, 173

Wakker, B. and Schwarz, U., 1988, Ann. Rev. Astron. Astrophys. 200, 312

Weir, N., 1991, in 3rd ESO/ST-ECF Data Analysis Workshop

Weir, N., 1992, in D. Worral, C. Biemesderfer, and J. Barnes (eds.), Astronomical Data Analysis Software and System 1, pp 186–190, Astronomical Society of the Pacific

Chapitre 5

Sélection d'articles publiés