



Université des Sciences et de la Technologie Houari Boumedienne
Institut d'Electronique

Thèse

Pour l'obtention du diplôme de

Magister en Electronique

Option : Electronique des Systèmes

ELIMINATION DU BRUIT DANS LES IMAGES PAR DECOMPOSITION EN ONDELETTES.

Présentée par : **Abdelhamid BOUDOUANI**

Soutenue le : 24 Mars 2001.

Devant le jury composé de :

Mr	AÏSSANI Amar Professeur (USTHB)	Président
Mme	BELHADJ AÏSSA Aïchouche Maître de conférences (USTHB)	Examinatrice
Mr	SMARA Youcef Maître de conférences (USTHB)	Examineur
Mr	EL BAHY Mostefa Chargé de cours (USTHB)	Examineur
Mr	HOUACINE Amrane Maître de conférences (USTHB)	Directeur de thèse

Remerciements

Je tiens à remercier Monsieur Houacine, mon Directeur de thèse, Maître de conférences à l'USTHB, pour m'avoir proposé le sujet de cette thèse, et je lui exprime toute ma reconnaissance pour tous les efforts qu'il a fourni jusqu'à l'aboutissement de ce travail.

Je tiens également à remercier le président du jury Monsieur Aïssani, Professeur à l'USTHB ainsi que les membres du jury Madame Belhadj Aïssa, Maître de conférences à l'USTHB, Monsieur Smara, Maître de conférences à l'USTHB, et Monsieur El Bahi Chargé de cours à l'USTHB, pour l'honneur qu'ils me font en acceptant d'évaluer ce travail

Je tiens aussi à remercier ma famille en particulier mes parents pour tous leurs encouragements, soutiens et sacrifices , je leurs exprime à cette occasion toute ma gratitude.

Table des Matières

Introduction générale	6
Chapitre 1 : Algorithme de la transformation en ondelettes	
1.1. La transformation de Fourier à court terme	8
1.2. Transformation en ondelettes continue	9
1.3. Analyse multirésolution et ondelettes	12
1.4. Algorithme de Mallat	14
1.5. Problème des bords	17
1.6. Ondelettes et moments	17
1.7. Propriétés fréquentielles de la fonction d'échelle et de l'ondelette	20
1.8. Bases biorthogonales	22
1.9. Passage à deux dimensions	25
1.10. Algorithme à trous	28
1.11. Transformation en ondelettes sans décimation	31
 Chapitre 2 : Seuillage des coefficients en ondelettes	 36
2.1 Réduction du bruit par seuillage des coefficients d'ondelette.	38
2.2. L'aspect non linéaire.	39
2.3. Erreur quadratique moyenne et fonction risque.	41
2.4. Propriété asymptotique du seuil à minimum de risque.	46
2.5. Le seuil universel.	
 Chapitre 3 : Estimation de l'erreur quadratique moyenne	 49
3.1. Estimateur SURE de l'EQM	50
3.2. SURE adaptative	51
3.3. Validation croisée ordinaire	52
3.4. Validation croisée généralisée	53
3.5. Comportement asymptotique	54
3.6. GCV pour un nombre de données fini	

3.7 Procédure de minimisation	54
Chapitre4 : Seuillage et GCV dans des situations plus réalistes	
4.1. Seuillage global d'un bruit coloré	58
4.2. Seuillage universel dépendant des niveaux	62
4.3. Extension au cas SURE et GCV	62
4.4. Transformation non orthogonale	64
4.5. Adaptivité pour chaque échelle	68
4.6. Transformation en ondelettes non décimée	68
Chapitre5 : Résultats et évaluations	
5.1. Evaluation qualitative	72
5.2. Evaluation quantitative	83
Conclusion générale	94
Bibliographie	95

Symboles et abréviations

$A_j(s)$: le projecteur de s sur l'espace V_j .

a : le facteur d'échelle de la TO.

b : le facteur de translation de la TO.

$B_{p,q}^s$: l'espace de Besov avec comme paramètres de douceur s , p et q .

Card : le cardinal d'un ensemble.

$C_s(a,b)$: la transformée en ondelettes du signal s au point b et à l'échelle a .

C_ψ : la constante de la condition d'admissibilité de l'ondelette ψ .

$D_j(s)$: le projecteur de s sur l'espace W_j .

db : abréviation de Décibel.

d_n^j : le détail à l'échelle j et à la position n .

$E(s)$: l'espérance mathématique du signal s .

EQM : l'erreur quadratique moyenne.

f : le filtre d'interpolation.

G : la fenêtre d'analyse dans la transformée de Gabor

h : le filtre passe-bas du banc de filtre utilisé pour la transformation en ondelettes.

$\tilde{h}$: le filtre h retourné tel que $\tilde{h}(n) = h(n)$

$\bar{h}$: le filtre dual de h .

g : le filtre passe-haut du banc de filtre utilisé pour la transformation en ondelettes.

$\tilde{g}$: le filtre g retourné tel que $\tilde{g}(n) = g(n)$

$\bar{g}$: le filtre dual de g .

GCV : la validation croisée généralisée.

I : la matrice identité.

L^2 : l'espace vectoriel des fonctions de carré intégrable.

M_ψ^k : le moment d'ordre k de l'ondelette ψ .

N : le nombre de donnée.

$\mathbb{R}$: l'ensemble des nombres réels.

OCV : la validation croisée ordinaire.

Q : la matrice de covariance du bruit.

s : le signal à analyser.

$\hat{s}$: la transformée de Fourier du signal s .

s_n^j : l'approximation à l'échelle j et à la position n du signal s .

S : la matrice de covariance des coefficients en ondelettes du bruit.

SURE : l'estimateur de risque non biaisé de Stein.

TF : la transformation de Fourier.

TG : la transformation de Gabor.

TO : la transformation en ondelettes.

x : la variable spatiale.

y : le signal bruité .

v : un coefficient d'ondelette non bruité.

V_j : l'espace vectoriel de l'échelle j .

w : un coefficient d'ondelette bruité.

w_λ : un coefficient d'ondelette après seuillage avec un seuil λ .

$\tilde{W}$: matrice de la transformation en ondelettes.

α : une mesure de préservation de contours.

ρ : la mesure de corrélation.

ε : l'entropie.

δ_{ij} : le symbole de kronecker.

$\Delta_{s(t)}$: la durée temporelle utile (résolution temporelle).

$\Delta_{\hat{s}(t)}$: la largeur utile du spectre (résolution fréquentielle).

ψ : l'ondelette analysante.

ψ^* : le complexe conjugué de ψ

φ : la fonction d'échelle.

ω, ξ : pulsation ($\omega=2\pi$ fréquence).

ω : transformation en ondelettes du bruit.

λ : le seuil appliqué aux coefficients d'ondelette.

λ^* : le seuil optimal.

λ_{univ} : le seuil universel.

ϕ : la distribution du bruit.

Φ : la fonction de répartition du bruit

σ : le niveau ou l'écart type du bruit

∇^2 : approximation du Laplacien.

Liste des figures et tableaux

– 1^{ier} Chapitre –

Fig 1.1 Représentation temps–fréquence et temps–échelle

Fig 1.2 Schémas de décomposition et de reconstruction de l'algorithme de Mallat avec décimation à une dimension.

Fig 1.3 Schémas de décomposition et de reconstruction de l'algorithme de Mallat sans décimation à une dimension.

Fig 1.4 Schémas de décomposition et de reconstruction de l'algorithme à trous .

Fig1.5 Schéma de décomposition et de reconstruction de l'algorithme de Mallat sans décimation à deux dimensions

Fig1.6 Schéma de décomposition et de reconstruction de l'algorithme de Mallat avec décimation à deux dimensions

Fig 1.7 Une image 256 niveaux de gris 'Léna'.

Fig 1.8 Représentation de la décomposition de l'image dans Fig 1.7 selon l'algorithme de Mallat.

Fig 1.9 Représentation de la décomposition de l'image dans Fig 1.7 selon l'algorithme à trous.

Fig 1.10 Représentation de la décomposition de l'image dans Fig 1.7 selon l'algorithme de Mallat sans décimation.

Fig 1.11 Un signal représentant une ligne de l'image dans Fig 5.2.

Fig 1.12 Transformation en ondelettes avec décimation du signal de la Fig 1.11.

Fig 1.13 Transformation en ondelettes sans décimation du signal de la Fig 1.11.

– 2^{ième} Chapitre –

Fig 2.1 Représentation des coefficients d'ondelette après seuillage en fonction des coefficients d'ondelette bruités.

Fig 2.2 Biais , variance et risque en fonction du seuil.

Fig 2.3 Contribution à la fonction risque en fonction du seuil pour différentes valeurs de coefficients d'ondelette.

Fig 2.4. Seuil optimal et son équivalent asymptotique en fonction du nombre de données N

– 3^{ème} Chapitre –

Fig 3.1 GCV et EQM de la fonction résultante en fonction du seuil

Fig 3.2 GCV et EQM de la fonction résultante en fonction du seuil pour $N=16384$ et $N=4096$.

Fig 3.2 GCV et EQM de la fonction résultante en fonction du seuil pour différents nombres de coefficients d'ondelette N .

Fig 3.3 Représentation de l'EQM et de son estimation GCV en fonction du seuil pour différents nombres de coefficients d'ondelette.

– 4^{ème} Chapitre –

Fig 4.1 Le signal original.

Fig 4.2 Le signal bruité avec un bruit coloré.

Fig 4.3 L'EQM de la fonction résultante en fonction du seuil.

Fig 4.4 TO du signal bruité avec un bruit coloré.

Fig 4.5 Le signal débruité avec un seuil qui minimise l'EQM.

Fig 4.6 GCV et EQM d'un signal bruité avec un bruit coloré.

Fig 4.7 GCV et EQM en fonction du seuil pour les niveaux 1, 2 et 3 (le seuil est différent d'un niveau à un autre), la transformation en ondelette est avec décimation.

Fig 4.8. Résultat du seuillage universel dépendant des niveaux.

Fig 4.9 Résultat du seuillage SURE dépendant des niveaux

Fig 4.10 Résultat du seuillage GCV dépendant des niveaux.

Fig 4.11 GCV et EQM en fonction du seuil pour les niveaux 1, 2 et 3 (le seuil est différent d'un niveau à un autre), la transformation en ondelette est sans décimation

Fig 4.12 . Résultat du seuillage universel dépendant des niveaux avec une transformation avec décimation

Fig 4.13 Résultat du seuillage SURE dépendant des niveaux avec une transformation sans décimation.

Fig 4.14 Résultat du seuillage GCV dépendant des niveaux avec une transformation sans décimation.

– 5^{ème} Chapitre –

Fig 5.1 Image originale ‘une maison’ (256 niveaux de gris).

Fig 5.2 Image bruitée (bruit blanc additif) à 10db.

Fig 5.3 Résultat du filtre médian appliqué à Fig5.2.

Fig 5.4 Résultat du filtre de Wiener appliqué à Fig 5.2.

Fig 5.5 Résultat du seuillage de Fig 5.2 avec un seuil universel global.

Fig 5.6 Résultat du seuillage de Fig 5.2 avec un seuil SURE global.

Fig 5.7 Résultat du seuillage de Fig 5.2 avec un seuil GCV global.

Fig 5.8 Résultat du seuillage de Fig 5.2 avec un seuil universel dépendant des niveaux.

Fig 5.9 Résultat du seuillage de Fig 5.2 avec un seuil SURE dépendant des niveaux.

Fig 5.10 Résultat du seuillage de Fig 5.2 avec un seuil GCV dépendant des niveaux.

Fig 5.11 Résultat du seuillage de Fig 5.2 avec un seuil universel dépendant des niveaux dans le cas où la transformation en ondelette est sans décimation.

Fig 5.12 Résultat du seuillage de Fig 5.2 avec un seuil SURE dépendant des niveaux dans le cas où la transformation en ondelette est sans décimation.

Fig 5.13 Résultat du seuillage de Fig 5.2 avec un seuil GCV dépendant des niveaux dans le cas où la transformation en ondelette est sans décimation.

Fig 5.14 Image bruitée avec un bruit coloré SNR=10db.

Fig 5.15 Résultat du filtre médian appliqué à Fig 5.14 (bruit coloré).

Fig 5.16 Résultat du filtre de Wiener appliqué à Fig 5.14 (bruit coloré).

Fig 5.17 Résultat du seuillage de Fig 5.14 avec un seuil universel global.

Fig 5.18 Résultat du seuillage de Fig 5.14 avec un seuil SURE global.

Fig 5.19 Résultat du seuillage de Fig 5.14 avec un seuil GCV global.

Fig 5.20 Résultat du seuillage de Fig 5.14 avec un seuil universel dépendant des niveaux.

Fig 5.21 Résultat du seuillage de Fig 5.14 avec un seuil SURE dépendant des niveaux.

Fig 5.22 Résultat du seuillage de Fig 5.14 avec un seuil GCV dépendant des niveaux.

Fig 5.23 Résultat du seuillage de Fig 5.14 avec un seuil universel dépendant des niveaux et où la transformation en ondelette est sans décimation.

Fig 5.24 Résultat du seuillage de Fig 5.14 avec un seuil SURE dépendant des niveaux et où la transformation en ondelette est sans décimation.

Fig 5.25 Résultat du seuillage de Fig 5.14 avec un seuil GCV dépendant des niveaux et où la transformation en ondelette est sans décimation.

Tab 5.1 Rapport signal sur bruit (Cas d'un seuillage global et d'un bruit blanc)

Tab 5.2 Rapport signal sur bruit (Cas d'un seuillage global et d'un bruit coloré)

Tab 5.3 Erreurs normalisées des images résultantes (Cas d'un seuillage global et d'un bruit blanc)

Tab 5.4 Erreurs normalisées des images résultantes (Cas d'un seuillage global et d'un bruit coloré)

Tab 5.5 Les mesures de corrélation et de préservation de contours (Cas d'un seuillage global et d'un bruit blanc)

Tab 5.6 Les mesures de corrélation et de préservation de contours (Cas d'un seuillage global et d'un bruit coloré)

Tab 5.7 Performances des filtres médian et de Wiener (Cas d'un seuillage global et d'un bruit blanc)

Tab 5.8 Performances des filtres médian et de Wiener (Cas d'un seuillage global et d'un bruit coloré)

Tab 5.9 Les performances avant seuillage (bruit blanc)

Tab 5.10 Performances après seuillage dépendant des niveaux (TO avec décimation)

Tab 5.11 Performance après seuillage dépendant des niveaux (TO sans décimation)

Tab 5.12 Les performances avant seuillage (bruit coloré)

Tab 5.13 Performances après seuillage (TO avec décimation)

Tab 5.14 Performances après seuillage (TO sans décimation)

Introduction générale

Bien que les techniques d'acquisition d'images soient en évolution permanente, la plupart des systèmes imageurs fournissent des images bruitées, ce qui rend leur exploitation directe difficile. Souvent, on recourt à des techniques post-acquisition pour l'élimination du bruit. Les techniques spatiales ont été les premières à être utilisées dans ce domaine, mais elles ont l'inconvénient de lisser les contours. D'où l'intérêt d'utiliser des techniques se basant sur des informations spectrales en plus des informations spatiales.

Par ailleurs une transformation espace-échelle est mieux adaptée aux signaux non-stationnaires comme les images. Une telle transformation se base sur la théorie des ondelettes et des algorithmes rapides ont été conçus permettant l'obtention de cette représentation d'une manière simple. Les algorithmes les plus connus dans ce domaine sont l'algorithme de Mallat qui permet d'obtenir une représentation multi-résolution, l'algorithme à trous sans décimation qui est invariant par translation et l'algorithme de Mallat sans décimation qui rassemble les avantages des algorithmes précédents.

Nous concevons l'élimination du bruit comme une opération d'estimation d'une fonction inconnue à partir de données bruitées avec comme principal objectif la minimisation de l'erreur quadratique moyenne. Plus précisément, nous proposons les techniques se basant sur le seuillage des coefficients d'ondelette. Pour estimer l'erreur quadratique moyenne nous avons utilisé l'estimateur de Stein non biaisé (SURE) cependant cet estimateur nécessite une mesure du niveau du bruit. Un autre estimateur qui est la validation croisée généralisée (GCV) ayant l'avantage de ne nécessiter aucune connaissance a priori sur le niveau du bruit a été utilisé.

Nous étendrons cette méthode d'élimination de bruit au cas d'images affectées par un bruit coloré, en effectuant un seuillage dépendant des niveaux et nous prospecterons l'apport d'une transformation invariante par translation à l'amélioration des résultats.

Pour évaluer les performances de ces techniques, nous considérons le rapport signal sur bruit des images débruitées. Nous complétons notre évaluation par des mesures de corrélation et de préservation de contours.

Cette thèse est organisée de la manière suivante : le premier chapitre est consacré à une présentation de la transformation en ondelettes et les algorithmes rapides pour son implémentation. Le deuxième chapitre présentera le principe du seuillage des coefficients en ondelettes et l'erreur quadratique moyenne comme critère d'élimination de bruit. Les estimateurs de l'erreur quadratique moyenne seront exposés dans le troisième chapitre. Dans le quatrième chapitre, nous examinerons l'extension de cette méthode de filtrage aux images corrompues par un bruit coloré. Enfin, le cinquième et dernier chapitre contiendra les résultats avec l'évaluation qualitative et quantitative.

1.1. La transformée de Fourier à court terme:

La transformée de Fourier à court terme, consiste en la projection d'un signal s sur un ensemble de fonctions bien localisées en temps et en fréquence (fenêtres). Elle est définie par le produit scalaire [Gab40]:

$$TG(s)(\omega, b) = \langle s, G_{\omega, b} \rangle = \int_{\mathbb{R}} s(x) G^*(x - b) e^{-j\omega x} dx \quad (1.1)$$

Où $G_{\omega, b}(x) = e^{j\omega x} G(x - b)$. $(\omega, b) \in \mathbb{R}^2$

La formule (1.1) s'exprime dans le domaine fréquentiel grâce au théorème de Plancherel :

$$TG(s)(\omega, b) = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{s}(\xi) \hat{G}(\xi - \omega) e^{-jb(\xi - \omega)} d\xi \quad (1.2)$$

Un certain nombre de fenêtres est utilisé telle que Hamming, Hanning ou Gaussienne. Dans ce dernier cas, la transformation est appelée Transformée de Gabor ‘‘ TG’’, où on déduit de l'écriture précédente de $G_{\omega, b}$, que celle-ci est une gaussienne modulée en amplitude par une fonction sinusoïdale de fréquence ω et décalée de la quantité b (en temps). $TG(s)(\omega, b)$ décrit en même temps le comportement temporel de $s(x)$ dans le voisinage d'un point $b = x$ et le comportement fréquentiel de $\hat{s}(\xi)$ dans le voisinage de $\omega = \xi$ d'où la dénomination analyse temps-fréquence.

La modification des valeurs de b et de ω permet un balayage de l'espace temps-fréquence. Celle-ci peut être représentée graphiquement par un spectrogramme.

La TG est une transformée locale qui ne permet de représenter que le contenu spectral d'une portion du signal localisée dans le temps. La transformation de Fourier étant très bien adaptée à l'analyse des signaux stationnaires (une description globale sans intervention du temps dans la TF est suffisante pour les signaux stationnaires), la TG sera donc très bien adaptée pour les signaux localement stationnaires (à l'intérieur d'un intervalle temporel égal à la largeur de la fenêtre G). Mais la validité de la TG est donc conditionnée par la taille de la fenêtre de l'analyse et par le signal que l'on désire analyser. Car si pour un signal donné, la fenêtre d'analyse est très étroite, la TG aura une résolution temporelle excellente, permettant de prendre en compte les non-stationnarités, mais les parties stationnaires du signal seront mal analysées du fait que la bande passante de la fenêtre (localisation fréquentielle) est trop large. Un cas extrême pour le choix de la fenêtre est une impulsion qui est parfaitement localisée en temps mais a une bande passante infinie. Si par contre la fenêtre est relativement large les parties non stationnaires (discontinuités) seront mal localisées, mais le contenu fréquentiel

local du signal sera très bien connu : là aussi un cas extrême de choix de fenêtre sera une sinusoïde de support temporel infini mais parfaitement localisée en fréquence (une raie).

Ceci pose un réel problème de choix de fenêtre, car on ne dispose pas d'a priori (sauf exception) sur les caractéristiques des signaux qu'on veut analyser. C'est notamment le cas des images qui comportent à la fois des non-stationnarités (contours), mais également des zones stationnaires (textures). Cependant, il est possible de construire des fenêtres de forme optimale qui réalisent un compromis entre les résolutions temporelle et fréquentielle tout en respectant une limitation fondamentale formulée dans le principe d'incertitude qui pour une paire de transformées $(s(t), \hat{s}(\omega))$ stipule que :

$$\Delta_{s(t)} \Delta_{\hat{s}(\omega)} \geq \frac{1}{2} \quad (1.3)$$

$$\text{où } \Delta_{s(t)}^2 = \frac{\int t^2 |s(t)|^2 dt}{\int |s(t)|^2 dt}$$

$$\Delta_{\hat{s}(\omega)}^2 = \frac{\int \omega^2 |\hat{s}(\omega)|^2 d\omega}{\int |\hat{s}(\omega)|^2 d\omega}$$

Donc pour une fenêtre de bande passante $\Delta_{\hat{s}(\omega)}$, sa longueur doit être plus grande que $\frac{1}{2\Delta_{\hat{s}(\omega)}}$

et vice versa.

La TG ne tient pas compte de la nature multi-échelle des images car elle suppose qu'elles sont composées de structures ayant chacune une localisation temporelle ou fréquentielle approximativement égale à la durée utile $\Delta_{s(t)}$ de la fenêtre d'analyse et à la bande passante $\Delta_{\hat{s}(\omega)}$.

le temps de calcul de la TG est très important dans le cas des images.

1.2. Transformée en ondelette continue

La première définition de la transformée en ondelette (TO) fut établie par Grossmann et Morlet en temps et échelle continue [Gro84]. Elle consiste à remplacer le paramètre de fréquence de la TG par un paramètre d'échelle. La transformée ainsi définie conduit à une analyse temps-échelle ou espace-échelle.

Définition

Soit $s \in L^2(\mathfrak{R})$ et $(a, b) \in \mathfrak{R}_+^* \times \mathfrak{R}$, l'analyse en ondelettes de la fonction s est définie par le produit scalaire :

$$C_s(a, b) = \langle s, \psi_{a,b} \rangle = a^{-1/2} \int_{\mathfrak{R}} s(x) \psi^* \left(\frac{x-b}{a} \right) dx \quad (1.4.a)$$

où $\psi^* \left(\frac{x-b}{a} \right)$ est le complexe conjugué de l'ondelette à l'échelle $\mathbf{a}$, centrée en $\mathbf{b}$.

La variable $\mathbf{a}$ joue le rôle inverse de la fréquence. On peut interpréter cette expression comme une projection du signal sur une famille de fonctions analysantes $\psi_{a,b}$, telle que :

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}} \psi \left(\frac{x-b}{a} \right)$$

La formule (1.4.a) s'exprime également dans le domaine de Fourier en utilisant le théorème de Plancherel par :

$$C_s(a, b) = \frac{a^{1/2}}{2\pi} \int_{\mathfrak{R}} \hat{s}(\xi) \hat{\psi}^*(a\xi) e^{i\xi b} d\xi \quad (1.4.b)$$

La reconstruction inverse de la fonction s à partir de sa TO continue est donnée par

$$s(x) = \frac{1}{C_\psi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} C_s(a, b) \psi_{a,b}(x) \frac{da}{a^2} db. \quad (1.5)$$

Où C_ψ est une constante donnée par $C_\psi = \int_{-\infty}^{+\infty} \frac{|\hat{\psi}(\xi)|^2}{|\xi|} d\xi$.

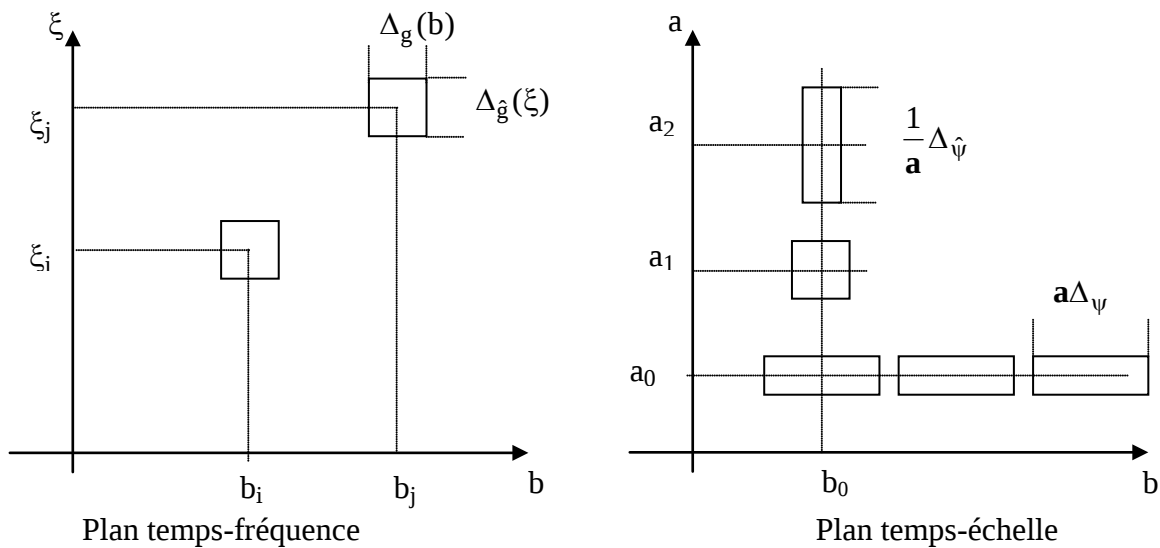


Fig 1.1 représentation temps-fréquence et temps-échelle

D'après la définition, on voit que pour reconstruire la fonction s il faut impérativement que la constante C_ψ soit finie. A partir de l'expression de cette constante on voit immédiatement qu'on a $C_\psi < \infty$ si et seulement si ψ est de moyenne nulle. (C'est à dire si ψ est intégrable et si $\hat{\psi}(0)=0$). Cette condition sur la fonction ψ est appelée condition d'admissibilité des ondelettes et a motivé le choix du vocable ondelette, car pour avoir une moyenne nulle il faut que ψ oscille. Si de plus ψ a une bonne décroissance alors elle oscillera sur une petite portion de l'axe des temps et ressemblera ainsi à une onde amortie. Une ondelette ψ qui convient doit avoir au moins une moyenne nulle et avoir un support le plus réduit possible [Dau88].

Le choix de l'ondelette est en principe ouvert il faut cependant noter que la convergence de l'algorithme de reconstruction donné par l'équation (1.5) est très dépendante de ce choix. La TO définie par (1.4.a) est isométrique : elle préserve l'énergie du signal analysé.

On remarque que $C_s(a, b)$ fournit une indication sur le comportement de $s(t)$ dans un voisinage du point $t = b$ (d'après 1.4 a) et de $\hat{s}(\xi)$ dans un voisinage du point $\xi = \frac{\xi_0}{a}$ (d'après 1.4.b), où ξ_0 est l'abscisse autour de laquelle $\hat{\psi}$ est concentrée, mais on ne peut pas pour autant qualifier $C_s(a, b)$ par une analyse temps–fréquence car le spectre du signal est analysé d'une manière indirecte, contrairement à la TG qui impose une fréquence d'analyse par la fréquence des oscillations de sa fonction analysante.

Comme pour la TG on note Δ_ψ et $\Delta_{\hat{\psi}}$, les durées utiles de l'ondelette ψ respectivement en temps et en fréquence. D'après ce que l'on vient de dire l'information sur le comportement de s analysée par $\psi_{a,b}$ sera connue avec la résolution “ $a \Delta_\psi$ ” dans le domaine temporel et la résolution “ $\frac{1}{a} \Delta_{\hat{\psi}}$ ” dans le domaine fréquentiel. On remarque que le produit “ $a \Delta_\psi \times \frac{1}{a} \Delta_{\hat{\psi}}$ ” est constant et indépendant de la valeur de l'échelle a . La partition du plan temps–échelle est donc réalisée par des cellules élémentaires d'aire constante (comme pour la TG mais de forme variable puisque les durées utiles de l'ondelette en temps et en fréquence varient, l'une par rapport à l'autre de façon inversement proportionnelle (l'une est multipliée par a , l'autre par a^{-1}). Cette différence fondamentale entre la TG et la TO continue fait que cette dernière est apte à analyser des signaux dans lesquels les événements se manifestent indistinctement à toutes les échelles.

La TO continue réalise une analyse à largeur de bande relative constante :

$$(Q = \frac{\text{fréquence centrale}}{\text{bande passante}} = \text{cte})$$

1.3 Analyse multirésolution et ondelettes

On considère l'espace vectoriel L^2 des fonctions de carré intégrable dans $\mathbb{R}$.

$$L^2 = \left\{ s : \int_{-\infty}^{+\infty} s^2(x) dx < \infty \right\}$$

Dans une analyse multirésolution on décompose l'espace L^2 en sous-espaces emboîtés V_j [Mal89].

$$\dots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset \dots$$

tel que le passage d'un espace à l'espace suivant se fait par changement d'échelle (Zoom), par exemple dans le cas dyadique on aura :

$$s(x) \in V_j \Leftrightarrow s(2x) \in V_{j-1}$$

Ce qui correspond à une dilatation d'un facteur de 2. L'espace V_j contient des signaux plus grossiers que l'espace V_{j-1} .

L'axiomatique correspondante peut être exprimée comme suit :

$$* \bigcup_{j=-\infty}^{+\infty} V_j = L^2.$$

$$* \bigcap_{j=-\infty}^{+\infty} V_j = \{0\}.$$

$$* s(x) \in V_j \Leftrightarrow s(2x) \in V_{j-1}$$

Dans ces conditions on peut montrer qu'il existe une fonction dite fonction d'échelle qui par dilatation et translation constitue une base pour les espaces V_j :

$$\exists \varphi(x) \in V_0 : \varphi_{j,k}(x) = 2^{-j/2} \varphi(2^{-j}x - k), \quad k \in \mathbb{Z} \quad (1.6)$$

φ est appelée la fonction mère.

La base sera orthonormée si :

$$\int_{-\infty}^{+\infty} \varphi(x) \varphi^*(x+n) dx = \delta(n), \quad \forall n \in \mathbb{Z}$$

L'action de projeter s fournit sa décomposition sur la base des fonctions d'échelles et les coefficients de cette décomposition constituent l'approximation à l'échelle j de s .

$$A_j = \sum_n \langle s, \varphi_{j,n} \rangle \varphi_{j,n}.$$

Les espaces V_j sont appelés espaces d'approximation.

Comme $V_j \subset V_{j-1}$, il y a des fonctions dans V_{j-1} qui n'appartiennent pas à V_j . Pour générer tous les éléments dans V_{j-1} à partir de V_j , on a besoin d'un espace supplémentaire.

On définit pour chaque espace V_j son complément orthogonal W_j dans V_{j-1}

Comme W_{j-1} est orthogonal à V_{j-1} , alors W_{j-1} sera orthogonal à W_j ($W_j \subset V_{j-1}$)

$$\forall j, k \neq j \quad W_j \perp W_k.$$

Les sous-espaces W_j ne forment pas une famille d'espaces emboîtés mais les propriétés d'échelle et d'invariance par translation sont conservées.

V_{j-1} peut ainsi être décomposé dans l'union des bases de V_j et W_j .

$$V_{j-1} = V_j \oplus W_j.$$

$$L^2(\mathbb{R}) = \bigoplus_{j \in \mathbb{Z}} W_j.$$

Dans ces conditions, on peut montrer qu'il existe une fonction ψ appelée fonction d'ondelette qui par dilatation et translation engendre une base orthonormée des W_j et donc de $L^2(\mathbb{R})$. Elle serait en plus l'ondelette mère, et pour tout sous-espace W_j , la base sera construite par la relation:

$$\psi_{j,k}(x) = 2^{-j/2} \psi(2^{-j}x - k).$$

L'approximation à l'échelle $j-1$ sera reconstruite à partir de l'approximation à l'échelle j , et en utilisant les détails du signal fournis par sa projection sur la base de W_j suivant la relation :

$$A_{j-1}(s) = A_j(s) + \sum_n \langle s, \psi_{j,n} \rangle \psi_{j,n}.$$

On notera $D_j(s)$ la projection de s dans W_j et le signal de détail sera décrit par

$$D_j(s) = \sum_n \langle s, \psi_{j,n} \rangle \psi_{j,n}.$$

Le signal détail est constitué d'une suite dont les éléments sont les coefficients en ondelette.

La formule de reconstruction s'écrit

$$A_{j-1}(s) = A_j(s) + D_j(s).$$

$$\langle \psi_{j,i}, \psi_{j',i'} \rangle = \delta_{j-j'} \delta_{i-i'}$$

en conséquence :

$$\forall k \in \mathbb{Z} : V_k \oplus \left(\bigoplus_{j=k}^{\infty} W_j \right) = L^2(\mathbb{R}).$$

Le problème pour effectuer correctement la décomposition est que l'on ne dispose pas en général du signal s mais seulement d'une approximation à une échelle donnée. Il faut donc trouver un algorithme qui à partir de cette approximation permet de trouver l'approximation et

les détails à la résolution immédiatement inférieure, ce qui par itération successive conduit à l'analyse du signal pour toutes les résolutions inférieures à celle de départ. Deux algorithmes principaux ont été mis en évidence, “ l'algorithme à trous ”[She92][Dja93] et “ l'algorithme de Mallat ”[Mal89], le premier concerne l'analyse non-orthogonale et le second est utilisé dans le cadre d'analyse multirésolution orthogonale et biorthogonale

1.4 Algorithme de Mallat

a. Etape de décomposition

Le point clef du développement de l'algorithme de Mallat [Mal89] est fourni par la décomposition de $s_n^{j-1} = \langle s, \varphi_{j-1,n} \rangle$ en fonction de $s_n^j = \langle s, \varphi_{j,n} \rangle$. Par construction $\varphi(x)$ est une fonction de V_0 . Comme V_0 est un sous-espace de V_{-1} : $V_0 \subset V_{-1}$; on peut exprimer $\varphi(x)$ en terme de la base de V_{-1} $\varphi_{-1}(x)$.

$$\varphi(x) = \sum_k h(k) \varphi_{-1}(x - k). \text{ soit } h(k) = \langle \varphi, \varphi_{-1} \rangle.$$

$$\varphi(x) = 2^{1/2} \sum_n h(k) \varphi(2x - k).$$

La suite $h(n)$ sera considérée comme étant la réponse impulsionnelle d'un filtre numérique dont la construction peut être faite à partir de la donnée de φ ou bien par un choix de $h(n)$ (certaines conditions sont à respecter), la fonction d'échelle est alors déterminée.

Montrons que la décomposition est la même pour des échelles quelconques, on a :

$$\varphi(x) = \sum_k h(k) 2^{1/2} \varphi(2x - k).$$

en remplaçant φ par sa nouvelle expression dans (1.6), on obtient:

$$\varphi_j(x) = 2^{-j/2} \sum_k h(k) 2^{1/2} \varphi(2(2^{-j}x - n) - k).$$

ce qui, en regroupant les indices et les exposants, conduit à l'équation

$$\varphi_{j,n}(x) = \sum_k h(k) \varphi_{j-1,k+2n}(x).$$

donc on peut calculer les coefficients $s_n^j = \langle s, \varphi_{j,n} \rangle$ à la résolution j par :

$$s_n^j(x) = \sum_k h(k) \langle s, \varphi_{j-1,k+2n} \rangle.$$

si on pose $l=k+2n$, cette équation s'écrit

$$s_n^j(x) = \sum_l h(l - 2n) \langle s, \varphi_{j-1,l} \rangle.$$

et si on note : $\tilde{h}(n) = h(-n)$. la séquence retournée ou le filtre symétrique de h , on obtient :

$$s_n^j(x) = \sum_l \tilde{h}(2n-l) \langle s, \varphi_{j-1,l} \rangle.$$

et on aura finalement l'équation réursive suivante :

$$s_n^j(x) = \sum_l \tilde{h}(2n-l) s_l^{j-1}.$$

Si on considère s_n^j comme une séquence numérique, le calcul précédent peut être interprété comme un produit de convolution entre $\tilde{h}$ et s^{j-1} évalué à un indice sur deux, soit un filtrage de la séquence s^{j-1} par un filtre de réponse impulsionnelle $\tilde{h}$ suivi par un sous-échantillonnage de rapport 2.

Du fait aussi que W_0 est un sous-espace de V_{-1} et dans l'objectif d'écrire $d_n^j = \langle s, \psi_{j,n} \rangle$ en fonction de $s_n^{j-1} = \langle s, \varphi_{j-1,n} \rangle$, un schéma analogue est bâti à partir de la décomposition de l'ondelette de W_0 sur la base de V_{-1} :

$$\psi = \sum_k g(k) \varphi_{-1,k}$$

ou de façon plus explicite :

$$\psi(x) = \sum_k g(k) 2^{1/2} \varphi(2x - k)$$

ce qui conduit à l'équation de construction de $g(n)$ suivante :

$$g(n) = \langle \psi, \varphi_{-1,n} \rangle.$$

$g(n)$ sera considérée comme la réponse impulsionnelle d'un filtre numérique. Nous verrons que le filtre $g(n)$ peut être construit à partir du filtre $h(n)$.

Un calcul analogue en tout point au précédent permet d'écrire les coefficients de détails :

$$d_n^j = \langle s, \psi_{j,n} \rangle,$$

$$d_n^j(x) = \sum_k g(k) \langle s, \psi_{j-1,k+2n} \rangle.$$

On introduit également le filtre symétrique dont la réponse impulsionnelle correspond à la séquence $g(n)$ retournée :

$$\tilde{g}(n) = g(-n).$$

la décomposition en ondelette à l'échelle j pourra donc s'écrire

$$d_n^j(x) = \sum_l \tilde{g}(2n-l) \langle s, \varphi_{j-1,l} \rangle.$$

ou encore

$$d_n^j(x) = \sum_l \tilde{g}(2n-l) s_l^{j-1}.$$

cette équation peut être interprétée de la même manière que précédemment.

b. Etape de reconstruction

La décomposition est gouvernée par l'équation

$$A_{j-1}(s) = \sum_n s_n^j \varphi_{j,n} + \sum_n d_n^j \psi_{j,n} \quad (1.7)$$

avec :

$$s_n^j = \langle s, \varphi_{j,n} \rangle \text{ et } d_n^j = \langle s, \psi_{j,n} \rangle$$

A_j est un projecteur donc $A_j(A_j(s)) = A_j(s)$. ce qui s'écrit :

$$A_{j-1}(s) = \sum_n \langle A_{j-1}(s), \varphi_{j-1,n} \rangle \varphi_{j-1,n} \quad (1.8)$$

En remplaçant dans (1.8) le terme $A_{j-1}(s)$ par son expression donnée par (1.7), on obtient l'équation suivante :

$$s_n^{j-1} = \sum_k s_k^j \langle \varphi_{j,k}, \varphi_{j-1,l} \rangle + \sum_k d_k^j \langle \psi_{j,k}, \varphi_{j-1,l} \rangle.$$

Or nous avons vu au paragraphe précédent que :

$$\varphi_{j,k} = \sum_l h(l) \varphi_{j-1,l+2k}$$

On peut donc évaluer le produit scalaire des séquences à deux résolutions successives quelconques par :

$$\langle \varphi_{j,k}, \varphi_{j-1,n} \rangle = \sum_l h(l) \langle \varphi_{j-1,l+2k}, \varphi_{j-1,n} \rangle = \sum_l h(l) \delta(n-l-2k).$$

ce qui donne :

$$\langle \varphi_{j,k}, \varphi_{j+1,n} \rangle = h(n-2k).$$

Et de même pour les ondelettes :

$$\langle \psi_{j,k}, \varphi_{j-1,n} \rangle = g(n-2k).$$

d'où l'équation de reconstruction :

$$s_n^{j-1} = \sum_k s_k^j h(n-2k) + \sum_k d_k^j g(n-2k).$$

Cette équation est une somme d'équations de filtrage (produit de convolution), si on remplace la suite s_k^j par la suite s_{2k}^j qui coïncide avec s_k^j pour $l=2k$: $s_{2k}^j = s_k^j$ et qui est nulle pour les valeurs intermédiaires (et de même pour la suite d_k^j).

$$s_k^j = \{\bullet, \bullet, \bullet, \bullet, \bullet\}$$

$$s_{2k}^j = \{\bullet, 0, \bullet, 0, \bullet, 0, \bullet, 0, \bullet\}$$

Cette opération consiste à intercaler un zéro entre les échantillons d'une série et correspond à un sur-échantillonnage. L'équation de reconstruction s'écrit alors :

$$s_n^{j-1} = \sum_k s_{2k}^j h(n-k) + \sum_k d_{2k}^j g(n-k).$$

1.5 Problème des bords :

Du fait que les signaux que nous traitons ont une taille finie, le calcul des coefficients en ondelette des parties se trouvant près des extrémités pose un problème car ce calcul nécessite la présence d'un voisinage de part et d'autre de l'échantillon considéré. Nous prolongeons ce signal au-delà de ces limites par symétrie.

Cette opération n'a pas une grande influence dans la suite des traitements dans le domaine des ondelettes car la zone qui nous intéresse dans le cas du traitement d'images est généralement loin des bords.

1.6 Ondelettes et Moments

Une des principales caractéristiques de la transformation en ondelette est de révéler l'aspect irrégulier des signaux [Fal94]. Pour cela, elle exhibe aux petites échelles des coefficients de grande amplitude là où le signal est irrégulier, et des coefficients nuls où quasi nuls là où il est régulier. L'insensibilité de la TO au comportement régulier est obtenue en imposant une condition d'orthogonalité entre les ondelettes et les polynômes. Cette condition est appelée condition de moments nuls ou d'oscillation. On note le moment d'ordre k de l'ondelette ψ par M_ψ^k , il est défini comme suit :

$$M_\psi^k = \langle x^k, \psi \rangle = \int_{\mathbb{R}} \psi(x) x^k dx.$$

On dit que l'ondelette ψ a p moments nuls si $M_\psi^k = 0$, pour $k=0, \dots, p-1$. Ce nombre est relié à la douceur de l'ondelette [Dau94].

Chaque moment nul impose des conditions sur ψ et sur le filtre g . l'utilisation d'un nombre de moments nuls élevé se répercute sur l'augmentation de la largeur de support de l'ondelette

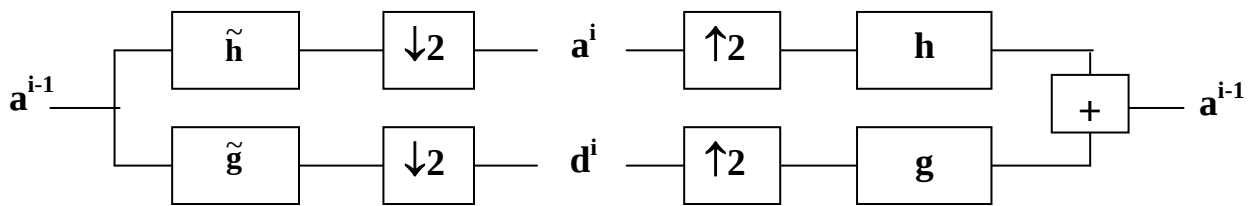


Fig 1.2 Schéma de décomposition et de reconstruction de l'algorithme de Mallat avec décimation

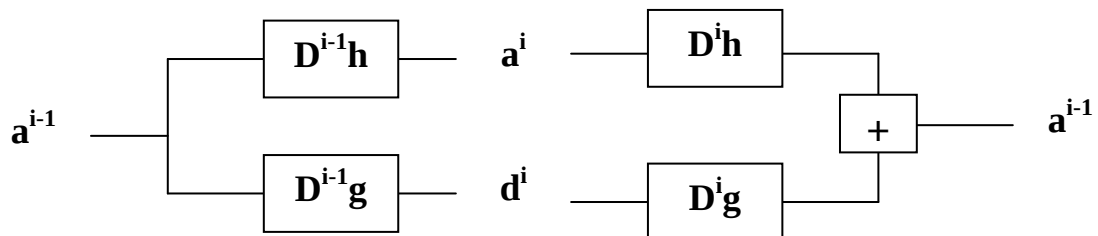


Fig 1.3 Schéma de décomposition et de reconstruction de l'algorithme de Mallat sans décimation

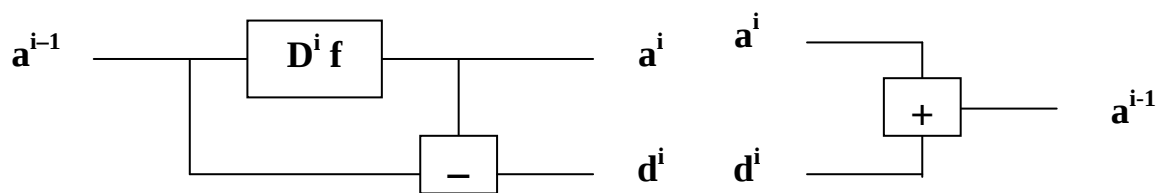


Fig 1.4 Schéma de décomposition et de reconstruction de l'algorithme à trous

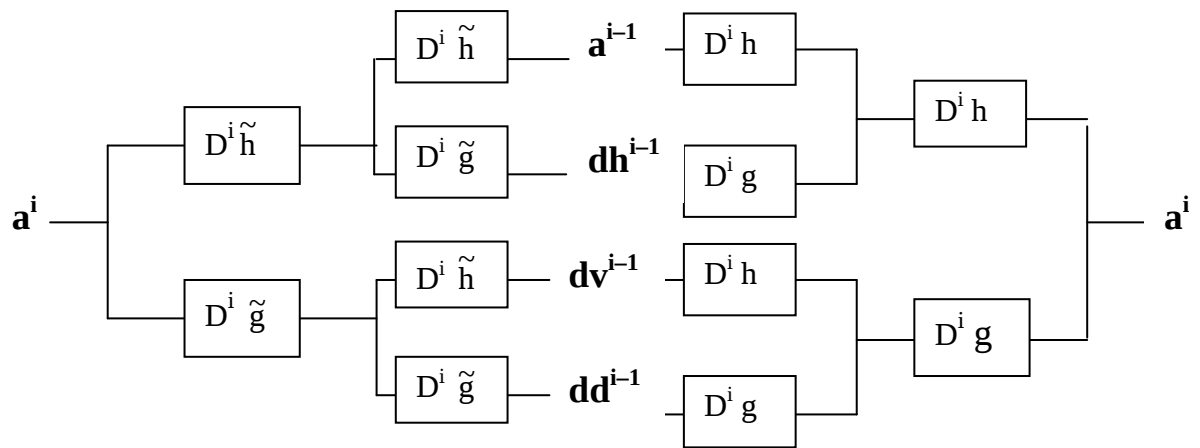


Fig 1.5 Schéma de décomposition et de reconstruction de l'algorithme de Mallat sans décimation à deux dimensions

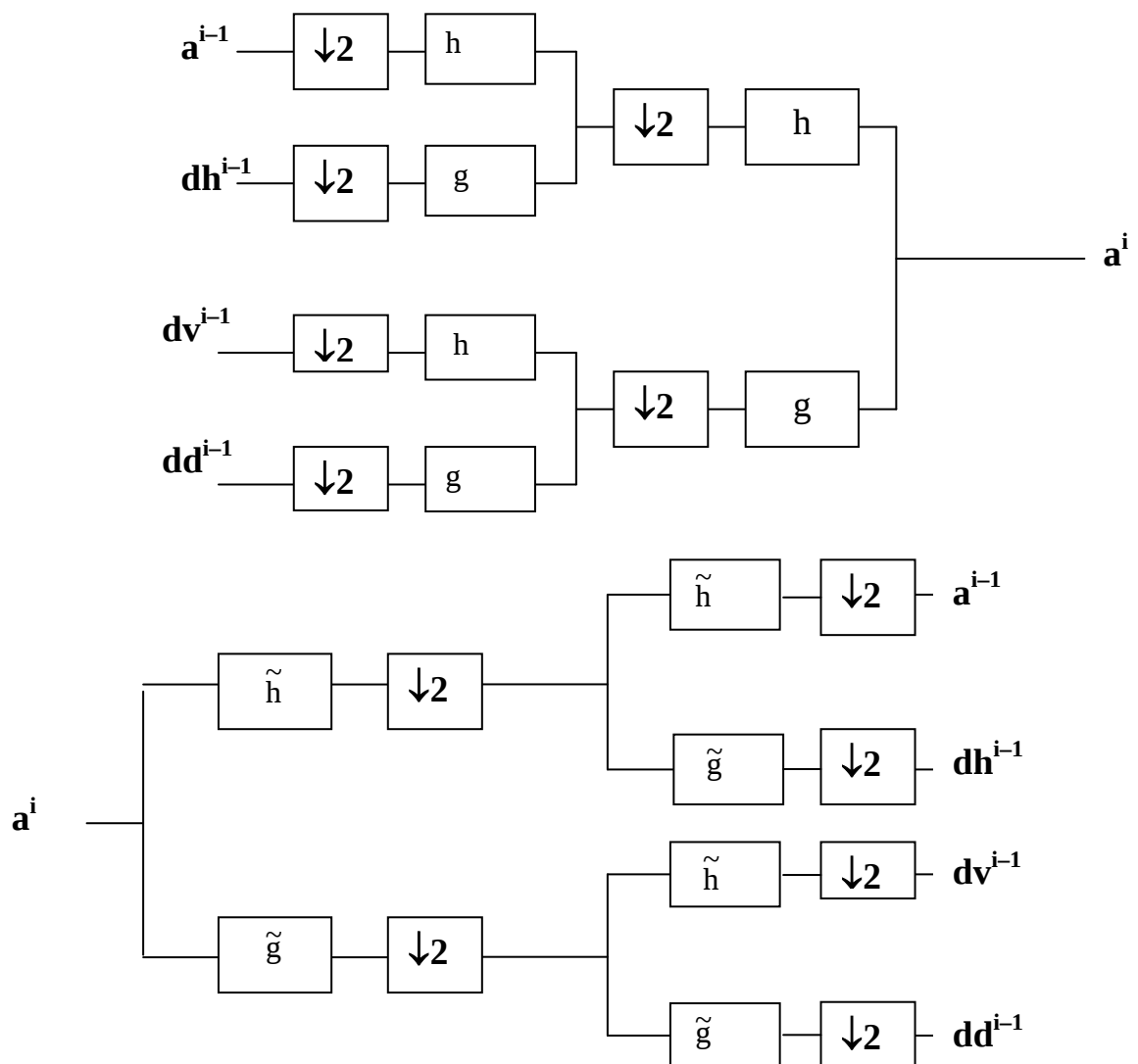


Fig 1.6 Schéma de décomposition et de reconstruction de l'algorithme de Mallat avec décimation à deux dimensions

et en conséquence sur la taille de filtres h et g . Pour cela on se limite dans le choix du nombre de moments nuls pour préserver la propriété de localisation.

1.7 Propriétés fréquentielles de la fonction d'échelle de l'ondelette

Pour construire ondelettes, fonctions d'échelle et filtres associés à une analyse multirésolution, on doit respecter certaines règles liées à l'orthogonalité des bases, nous verrons que la construction commence en général par le choix de la fonction d'échelle ou le filtre numérique associé, l'ondelette et le filtre associé peuvent alors être déduits. La raison de ce cheminement est que les contraintes pesant sur la fonction d'échelle sont moins lourdes que sur l'ondelette, (orthogonalité dans une même échelle seulement pour la fonction d'échelle, alors que l'ondelette doit être en plus orthogonale à travers les résolutions).

a. Propriétés fréquentielles de la fonction d'échelle :

La suite numérique $h(n)$ étant la réponse impulsionnelle d'un filtre numérique, soit $\hat{h}(\omega)$ sa transformée de Fourier.

$$\hat{h}(\omega) = \sum_n h(n)e^{-in\omega}$$

La transformée de Fourier de la fonction d'échelle est donnée par :

$$\hat{\phi}(\omega) = \frac{1}{\sqrt{2}} \hat{h}\left(\frac{\omega}{2}\right) \hat{\phi}\left(\frac{\omega}{2}\right).$$

Cette relation exprime la projection de la fonction d'échelle sur la résolution inférieure dans l'espace de Fourier. Par ailleurs la condition d'orthonormalité sur la fonction d'échelle s'exprime dans le domaine fréquentiel par :

$$\sum_n |\hat{\phi}(\omega + 2\pi n)|^2 = 1 \quad (1.9)$$

Cette relation est fondamentale si le choix de la fonction d'échelle se fait dans le domaine de Fourier (elle assure l'orthonormalité). On peut utiliser les deux derniers résultats pour arriver à la propriété suivante qui est fondamentale pour le filtre h :

$$|\hat{h}(\omega)|^2 + |\hat{h}(\omega + \pi)|^2 = 2. \quad (1.10)$$

Ce résultat nous permet de préciser la signification physique du filtre $h(n)$.

En admettant que la fonction d'échelle est de moyenne non nulle, on a :

$$\hat{h}\left(\frac{\omega}{2}\right) = \frac{\hat{\phi}(\omega)}{\hat{\phi}\left(\frac{\omega}{2}\right)}.$$

En appliquant ce résultat en zéro puis en π avec (1.10) nous obtenons :

$$\begin{cases} \hat{h}(0) = \sqrt{2}. \\ \hat{h}(\pi) = 0. \end{cases}$$

On pourra donc conclure que le filtre h est un filtre passe-bas ce qui correspond à l'idée intuitive que l'on a d'un passage d'une résolution à une résolution plus grossière en éliminant les détails. En utilisant ce dernier résultat et en utilisant la relation (1.9), on peut prouver que :

$$\hat{\phi}(0) = \int_{-\infty}^{+\infty} \phi(x) dx = 1.$$

b. Propriétés fréquentielles de l'ondelette :

Nous allons examiner, pour la fonction d'ondelette $\psi(x)$ et son filtre numérique associé $g(n)$, les propriétés équivalentes à celles trouvées précédemment. Ces propriétés devraient être plus contraignantes car les fonctions d'ondelette forment une base orthonormée à travers les résolutions et à l'intérieur d'une même résolution. La relation qui traduit la décomposition d'une ondelette dans la base de la fonction d'échelle à la résolution inférieure, s'exprime dans le domaine de Fourier par :

$$\hat{\psi}(\omega) = \frac{1}{\sqrt{2}} \hat{g}\left(\frac{\omega}{2}\right) \hat{\psi}\left(\frac{\omega}{2}\right).$$

L'orthonormalité d'une ondelette à une résolution conduit à la relation :

$$\sum_n |\hat{\psi}(\omega + 2\pi n)|^2 = 1$$

ce qui implique pour le filtre $g(n)$ la propriété suivante :

$$|\hat{g}(\omega)|^2 + |\hat{g}(\omega + \pi)|^2 = 2$$

Pour construire le filtre g , il faut tenir compte de l'orthogonalité entre ψ et ϕ qui s'exprime dans le domaine de Fourier par la relation suivante :

$$\sum_n \hat{\psi}(\omega + 2\pi n) \hat{\phi}(\omega + 2\pi n) = 0.$$

En appliquant cette relation en $\omega=0$ et en utilisant les propriétés on obtient :

$$\hat{\psi}(0) = \int_{-\infty}^{+\infty} \psi(x) dx = 0$$

contrairement à la fonction d'échelle la fonction d'ondelette à une moyenne nulle ce qui en fait d'ailleurs une ondelette admissible. Ceci se traduit pour le filtrage par :

$$\begin{cases} \hat{g}(0) = \sum_n g(n) = 0. \\ |\hat{g}(\pi)| = 1. \end{cases}$$

se qui signifie que g est un filtre passe-haut.

On montre qu'on peut écrire :

$$\hat{g}(\omega) = -\hat{\lambda}(\omega)\hat{h}^*(\omega + \pi)$$

avec :

$$\begin{cases} \hat{\lambda}(\omega) \text{ } 2\pi \text{ périodique.} \\ \hat{\lambda}(\omega) + \hat{\lambda}(\omega + \pi) = 0. \\ |\hat{\lambda}(\omega)| = 1. \end{cases}$$

La fonction $\hat{\lambda}(\omega)$ doit être choisie dans le cercle unité et doit respecter certaines conditions de symétrie, le choix est loin d'être unique mais on s'en tient au cas où la linéarité en phase serait préservée , le choix le plus simple est $\hat{\lambda}(\omega) = e^{-i\omega}$, ce qui donne :

$$\hat{g}(\omega) = -e^{-i\omega}\hat{h}^*(\omega + \pi)$$

ce qui, en effectuant la transformée de Fourier inverse, devient:

$$g(n) = (-1)^n h(1-n).$$

1.8 Bases Biothogonales

Malgré ses avantages, l'analyse multirésolution présente un inconvénient : si on veut préserver la linéarité en phase (symétrie droite gauche), on doit utiliser des filtres à réponse impulsionnelle infinie. Cependant la mise en œuvre de tels filtres est relativement lourde. Dans les applications en traitement d'image non seulement cette symétrie du traitement est très souhaitable mais en plus ce type d'application est particulièrement sensible à la complexité de l'algorithme du fait du nombre élevé de données à traiter. Ces deux éléments ont conduit les chercheurs à proposer une variante de l'analyse multirésolution conciliant linéarité en phase et compacité des supports des filtres associés en relaxant la propriété d'orthogonalité en faisant appel à des bases biorthogonales [Coh90]. Ces familles de fonctions assurent une analyse multirésolution du signal et assurent la reconstruction parfaite. L'analyse reste non-redondante, possède les propriétés de symétrie et de simplicité requise, mais conduit à une corrélation entre les composants du signal plus grande que dans le cas des bases

orthonormées. En plus, l'analyse et la reconstruction n'utilisent pas les mêmes familles de fonctions : deux jeux de bases duales sont nécessaires, ces familles sont orthogonales entre elles mais ne sont pas orthogonales en elles-mêmes.

Nous avons besoin de deux familles duales de sous-espaces emboîtés V_j et $\bar{V}_j$ définies par deux fonctions d'échelles, chacune formant une base non-orthogonale.

$$\begin{cases} V_j \subset V_{j-1} \\ \bar{V}_j \subset \bar{V}_{j-1} \end{cases}$$

Les sous-espaces complémentaires sont définis différemment car ils sont complémentaires aux sous-espaces duaux :

$$\begin{cases} W_j \perp \bar{V}_j \\ \bar{W}_j \perp V_{j-1} \end{cases}$$

$$V_{j-1} = V_j \oplus W_j \text{ avec } V_j \angle W_j$$

$$\bar{V}_{j-1} = \bar{V}_j \oplus \bar{W}_j \text{ avec } \bar{V}_j \angle \bar{W}_j$$

$$\text{Cependant } \bar{W}_j \perp W_j$$

Les propriétés d'orthogonalité vis-à-vis de la translation ou de la dilatation n'est requise que d'un espace à son dual.

$$\langle \psi_{j,n}, \bar{\psi}_{j,n} \rangle = \delta(i-j)\delta(m-n)$$

$$\langle \varphi_{j,n}, \bar{\varphi}_{j,n} \rangle = \delta(m-n)$$

L'analyse et la synthèse d'un signal se font d'une manière analogue à celle du cas orthonormal.

$$A_{j-1}(s) = A_j(s) + D_j(s).$$

$$A_j(s) = \sum_n \langle s, \bar{\varphi}_{j,n} \rangle \varphi_{j,n}$$

$$D_j(s) = \sum_n \langle s, \bar{\psi}_{j,n} \rangle \psi_{j,n}.$$

Les algorithmes récursifs seront identiques, avec pour l'analyse :

$$s_n^j = \sum_k \tilde{h}(2n-k) s_k^{j-1}.$$

$$d_n^j = \sum_k \tilde{g}(2n-k) s_k^{j-1}.$$

et pour la synthèse :

$$s_n^{j-1} = \sum_k s_k^j h(n-2k) + \sum_k d_k^j g(n-2k).$$

a. Propriété des bases et des filtres

Les filtres associés à chaque base de fonction d'échelle ont les mêmes propriétés que dans le cas orthonormal.

$$\hat{\phi}(\omega) = \frac{1}{\sqrt{2}} \hat{h}\left(\frac{\omega}{2}\right) \hat{\phi}\left(\frac{\omega}{2}\right).$$

$$\hat{\bar{\phi}}(\omega) = \frac{1}{\sqrt{2}} \hat{h}\left(\frac{\omega}{2}\right) \hat{\bar{\phi}}\left(\frac{\omega}{2}\right).$$

La relation d'orthogonalité des fonctions d'échelle prend une forme un peu différente :

$$\sum_n \bar{\phi}(\omega + 2\pi n) \hat{\phi}^*(\omega + 2\pi n) = 1$$

Cette relation permet de trouver une base duale si une fonction d'échelle est choisie. On obtient aussi une relation entre les filtres associés aux deux bases duales :

$$\hat{h}(\omega) \hat{h}^*(\omega) + \hat{h}(\omega + \pi) \hat{h}^*(\omega + \pi) = 2.$$

on peut vérifier que les deux filtres obtenus sont de type passe-bas et qu'ils ont les mêmes propriétés en zéro et π que dans le cas orthonormal :

$$\hat{h}(0) = \hat{h}(0) = \sqrt{2}.$$

$$\hat{h}(\pi) = \hat{h}(\pi) = 0.$$

Pour les fonctions d'ondelettes, on a :

$$\hat{\psi}(\omega) = \frac{1}{\sqrt{2}} \hat{g}\left(\frac{\omega}{2}\right) \hat{\phi}\left(\frac{\omega}{2}\right).$$

$$\hat{\bar{\psi}}(\omega) = \frac{1}{\sqrt{2}} \hat{g}\left(\frac{\omega}{2}\right) \hat{\bar{\phi}}\left(\frac{\omega}{2}\right).$$

L'orthonormalité vis-à-vis de la translation n'est vérifiée qu'entre une base et sa duale :

$$\sum_n \hat{\psi}(\omega + 2\pi n) \hat{\bar{\psi}}^*(\omega + 2\pi n) = 1.$$

ce qui conduit pour les filtres g associés à :

$$\hat{g}(\omega) \hat{g}^*(\omega) + \hat{g}(\omega + \pi) \hat{g}^*(\omega + \pi) = 2.$$

là aussi on peut vérifier que les propriétés des filtres en zéro et π restent inchangées, soit :

$$\hat{g}(0) = \hat{g}(0) = 0.$$

$$\hat{g}(\pi) = \hat{g}(\pi) = \sqrt{2}.$$

ce qui nous conduit à déduire que les deux filtres duaux sont de type passe-haut.

La même méthode que pour la fonction échelle permet de construire la base d'ondelette duale quand on connaît la base initiale. Il reste à relier la base initiale à la base de la fonction d'échelle choisie. Pour cela il faut exploiter l'orthogonalité entre l'ondelette avec le dual de la fonction d'échelle qui se traduit dans l'espace de Fourier par :

$$\sum_n \hat{\psi}(\omega + 2\pi n) \hat{\varphi}^*(\omega + 2\pi n) = 0.$$

D'où pour les filtres numériques associés :

$$\hat{g}(\omega) \hat{h}^*(\omega) + \hat{g}(\omega + \pi) \hat{h}^*(\omega + \pi) = 0.$$

à partir de cette relation, on peut montrer qu'une solution pour le choix du filtre g qui respecte la linéarité en phase est donnée par :

$$\hat{g}(\omega) = -e^{-i\omega} \hat{h}^*(\omega + \pi).$$

par transformée de Fourier inverse, on a :

$$g(n) = (-1)^n \bar{h}(1-n).$$

Nous avons ainsi les principaux éléments qui permettent de construire une analyse sur bases biorthogonales. Un choix intéressant pour une telle base est la famille des B-splines. Ces familles présentent les qualités de symétrie recherchées ainsi que la compacité des supports, propices à des mises en œuvre économiques.

1.9 Passage à deux dimensions

Les principes de l'analyse multirésolution sont facilement étendus aux signaux bidimensionnels (images), les fonctions que nous voulons traiter seront de carré intégrable sur $\mathcal{R}^2$ (les signaux appartiennent à $L^2(\mathcal{R}^2)$).

Nous considérons le cas particulier des bases séparables, où nous nous limiterons au cas dyadique.

La famille de fonctions d'échelle dilatées-translatées sera notée :

$$\varphi_{j,n,m}(x,y) = 4^{-j/2} \varphi(2^{-j}x-n, 2^{-j}y-m).$$

L'opération de dilatation se fait d'une manière indépendante suivant les deux directions, on crée une analyse multirésolution en choisissant comme fonction d'échelle le produit de deux fonctions d'échelle à une variable construites selon les règles déjà établies dans le cas 1D. Pour des raisons de simplicité et d'isotropie de traitement, on prendra deux fois la même fonction d'échelle :

$$\varphi_{j,n,m}(x,y) = 2^{-j/2} \varphi(2^{-j}x-n) 2^{-j/2} \varphi(2^{-j}y-m).$$

soit :

$$\varphi_{j,n,m}(x,y) = \varphi_{j,n}(x) \varphi_{j,m}(y).$$

la famille de sous-espaces ainsi construite est le produit tensoriel des sous-espaces à une dimension suivant les directions x et y :

$$V_{j-1}^2 = V_{j-1} \otimes V_{j-1} = (V_j \oplus W_j) \otimes (V_j \oplus W_j).$$

En développant :

$$V_{j-1}^2 = (V_j \otimes V_j) \oplus [(V_j \otimes W_j) \oplus (W_j \otimes V_j) \oplus (W_j \otimes W_j)].$$

Ce qui s'écrit également :

$$V_{j-1}^2 = V_j^2 \oplus W_j^2.$$

On constate que la base du sous-espace de détail sera constituée par trois familles d'ondelettes ($\psi^H(x,y) = \varphi(x)\psi(y)$, $\psi^V(x,y) = \psi(x)\varphi(y)$, $\psi^D(x,y) = \psi(x)\psi(y)$).

On remarque que dans les trois cas, les ondelettes sont séparables (comme la fonction d'échelle) et l'algorithme de transformation se déduit directement des algorithmes présentés pour le cas 1D. Les opérations de filtrage seront menées en cascade suivant les lignes puis les colonnes (ou inversement). Le signal 2D sera traité comme un signal 1D et l'analyse 1D est ainsi transposée à deux dimensions. La décomposition d'une image discrète considérée comme l'approximation de départ de l'image s^{j-1} fournit à chaque étape quatre nouvelles images quatre fois plus petites. L'une correspond à l'approximation inférieure s^j et les trois autres constituent les coefficients d'ondelette et contiennent les détails perdus lors du passage d'une résolution à la suivante. Ces détails sont orientés horizontalement d_h^j (correspondant à $\psi^H(x,y)$), verticalement d_v^j (correspondant à $\psi^V(x,y)$) ou diagonalement d_d^j (correspondant à $\psi^D(x,y)$).

En suivant le raisonnement précédent, dans le cas de l'algorithme de Mallat bidimensionnel quatre combinaisons se présentent donnant quatre résultats :

$$\text{L'image approximation au niveau suivant : } s_{n,m}^j = \sum_k \sum_l \tilde{h}(2n-k) \tilde{h}(2m-l) s_{kl}^{j-1}.$$

$$\text{Les détails horizontaux au niveau suivant : } dh_{n,m}^j = \sum_k \sum_l \tilde{h}(2n-k) \tilde{g}(2m-l) s_{kl}^{j-1}.$$

$$\text{Les détails verticaux au niveau suivant : } dv_{n,m}^j = \sum_k \sum_l \tilde{g}(2n-k) \tilde{h}(2m-l) s_{kl}^{j-1}.$$



Fig 1.7 Image 256 niveau de gris (léna)

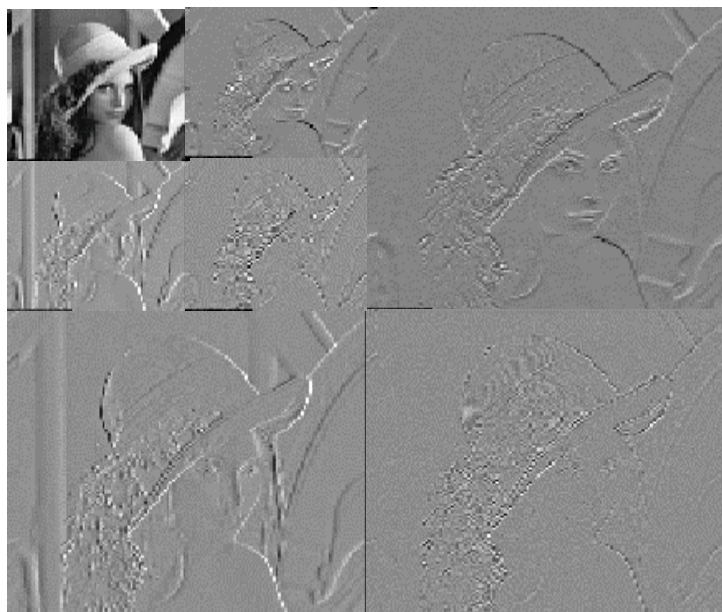


Fig 1.8 représentation de la décomposition en ondelettes (jusqu'à l'échelle 2) selon l'algorithme de Mallat de l'image dans la figure 1.7 :
on a les trois images détails à chaque échelle plus l'image approximation à la dernière échelle.

Les détails diagonaux au niveau suivant : $dd_{n,m}^j = \sum_k \sum_l \tilde{g}(2n-k) \tilde{g}(2m-l) s_{k,l}^{j-1}$.

La reconstruction se fait, dans ce cas, selon l'équation suivante :

$$s_{n,m}^{j-1} = \sum_k \sum_l s_{k,l}^j h(n-2k)h(m-2l) + \sum_k \sum_l dv_{k,l}^j h(n-2k)g(m-2l) \\ + \sum_k \sum_l dh_{k,l}^j g(n-2k)h(m-2l) + \sum_k \sum_l dd_{k,l}^j g(n-2k)g(m-2l);$$

On remarque que la surface occupée par les quatre images ensemble à la résolution j est la même que la surface occupée par l'image à la résolution $j-1$; il n'y a ni compression ni redondance dans la représentation.

Ondelette non-séparable :

L'analyse séparable dyadique nécessite trois familles d'ondelette, ceci est parfois considéré comme un inconvénient. D'autre part, le facteur de dilatation entre deux échelles successives est de 4 ce qui peut sembler élevé. Il est possible d'y remédier aux prix de la perte de la séparabilité des filtres, ce qui implique une complexité algorithmique plus élevée.

1.10 Algorithme à trous :

On considère l'algorithme à trous comme un banc de filtres à deux filtres f et g [She92] [Dja93], le filtre f qui est passe-bas est associé à une base d'interpolation $\varphi(x)$ tandis que le filtre g qui est passe haut est associé à l'ondelette $\psi(x)$, tel que :

$$\frac{1}{2^j} \varphi\left(\frac{x}{2^j}\right) = \frac{1}{2^{j-1}} \sum_n f(n) \varphi\left(\frac{x}{2^{j-1}} - n\right), \\ \frac{1}{2^j} \psi\left(\frac{x}{2^j}\right) = \frac{1}{2^{j-1}} \sum_n g(n) \psi\left(\frac{x}{2^{j-1}} - n\right),$$

On doit remarquer que f et g n'ont pas à satisfaire les contraintes précédemment citées dans l'algorithme de Mallat. En pratique le filtre f est un filtre interpolation B-spline.

L'approximation et les détails à l'échelle j sont obtenus en effectuant le filtrage de l'approximation à l'échelle précédente $j-1$ avec respectivement les filtres f et g ; le filtrage doit être précédé par une insertion de zéros entre les éléments du filtre créant des trous (d'où la dénomination), le pas entre éléments est 2^{j-1} . Ceci est résumé dans les équations suivantes :

L'approximation à l'échelle j :

$$s^j(n) = \sum_k \bar{f}(k) s^{j-1}(n + 2^{j-1}k),$$

d'où

$$s^j(n) = \sum_k \bar{f}\left(\frac{n-k}{2^{j-1}}\right) s^{j-1}(k),$$

Les détails à l'échelle j :

$$d^j(n) = \sum_k \bar{g}(k) s^{j-1}(n + 2^{j-1}k),$$

$$d^j(n) = \sum_k \bar{g}\left(\frac{n-k}{2^{j-1}}\right) s^{j-1}(k),$$

l'information perdue lors du passage entre deux approximations successives s^{j-1} et s^j représentent les coefficients en ondelettes d^j , ils sont déduits en effectuant la différence entre les images d'approximation :

$$d^j(n) = s^{j-1}(n) - s^j(n),$$

Cette déduction permet de simplifier l'algorithme, aussi bien à l'étape de décomposition qu'à l'étape de reconstruction. Celle-ci se fait à présent par simple addition des coefficients d'ondelettes aux différentes échelles avec la dernière approximation, soit :

$$s^0(n) = \sum_{i=1}^J d^i(n) + s^J(n)$$

où J est l'indice de la dernière échelle de décomposition et s^0 correspond au signal de départ.

La principale différence entre l'algorithme à trous et l'algorithme de Mallat est que les étapes de sous-échantillonnage dans l'analyse et sur-échantillonnage dans la synthèse sont supprimées et ont été substituées par une insertion de $(2^{j-1}-1)$ zéros entre les coefficients des filtres ($j-1$ est l'indice de l'échelle courante), la transformé devient sans décimation. Cette dernière opération donne à l'algorithme la propriété d'invariance par translation [She92][Nas93].

Dans le cas de l'algorithme à trous bidimensionnel, l'approximation à l'échelle suivante est obtenue selon l'équation :



Fig 1.9 représentation de la transformation en ondelettes de l'image de la figure 1.7 selon l'algorithme à trous

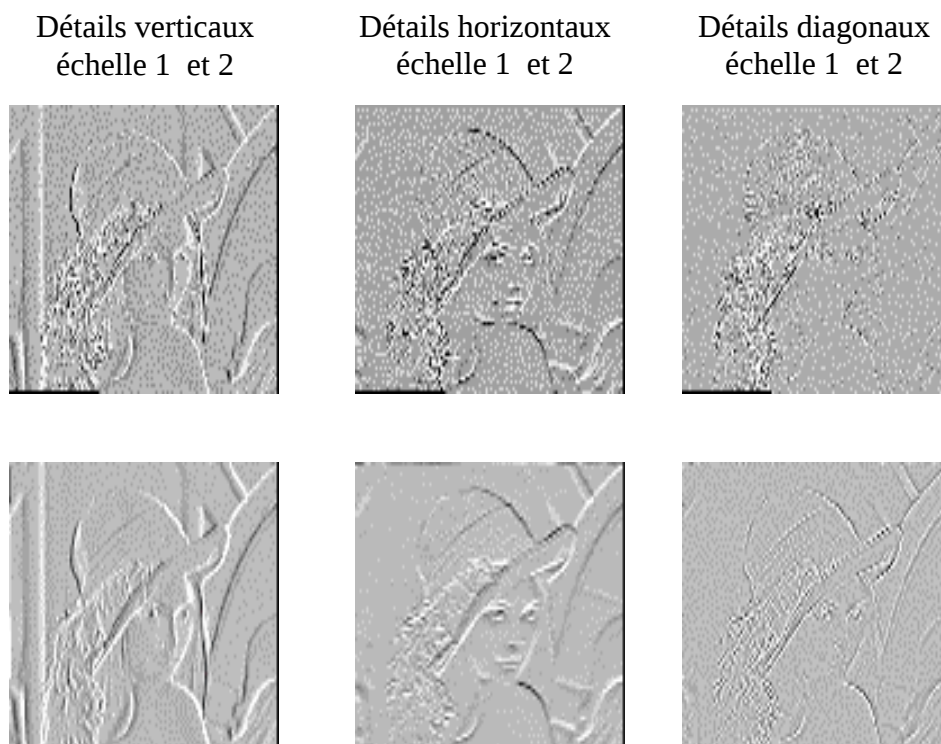


Image approximation à l'échelle 2



Fig 1.10 représentation de la transformation en ondelettes de l'image de la figure 1.7 selon l'algorithme de Mallat sans décimation .

$$s^j(n, m) = \sum_k \sum_l \bar{f}\left(\frac{n-k}{2^{j-1}}\right) \bar{f}\left(\frac{m-l}{2^{j-1}}\right) s^{j-1}(k, l);$$

les détails sont obtenus, de la même manière que dans le cas 1D, en effectuant la différence entre s^j et s^{j-1} . Dans ce cas les détails seront rassemblés dans un seul résultat :

$$d^j(n, m) = s^{j-1}(n, m) - s^j(n, m);$$

la reconstruction se fait simplement par addition des images détails avec l'approximation :

$$s^{j-1}(n, m) = d^j(n, m) + s^j(n, m);$$

1.11 Transformation en Ondelettes sans décimation :

Il a été démontré [She92] que la propriété d'invariance par translation de l'algorithme à trous vient du fait de la non-décimation au passage d'une échelle à la suivante. Pour généraliser cette propriété à toute transformation en ondelette une variante de l'algorithme de Mallat a été définie où on combine l'utilisation de filtres h et g qui réalisent les propriétés d'orthogonalité ou de biorthogonalité avec la non-décimation. Comme pour l'algorithme à trous la non-décimation engendre obligatoirement une dilatation des filtres par insertion de $2^{j-1}-1$ zéros entre les coefficients.

Equations de décomposition :

$$s^j(n) = \sum_k \tilde{h}\left(\frac{n-k}{2^{j-1}}\right) s^{j-1}(k);$$

$$d^j(n) = \sum_k \tilde{g}\left(\frac{n-k}{2^{j-1}}\right) s^{j-1}(k);$$

Equation de reconstruction :

$$s^{j-1}(n) = \sum_k h\left(\frac{n-k}{2^{j-1}}\right) s^j(k) + \sum_k g\left(\frac{n-k}{2^{j-1}}\right) d^j(k).$$

Nason [Nas96] a démontré que cette décomposition invariante par translation (stationnaire), contient les coefficients d'ondelettes décimés du signal original après avoir effectué une translation avec différents pas. Ce qui explique qu'à la reconstruction, le signal obtenu est équivalent à la moyenne des reconstructions obtenues avec une TO décimée.

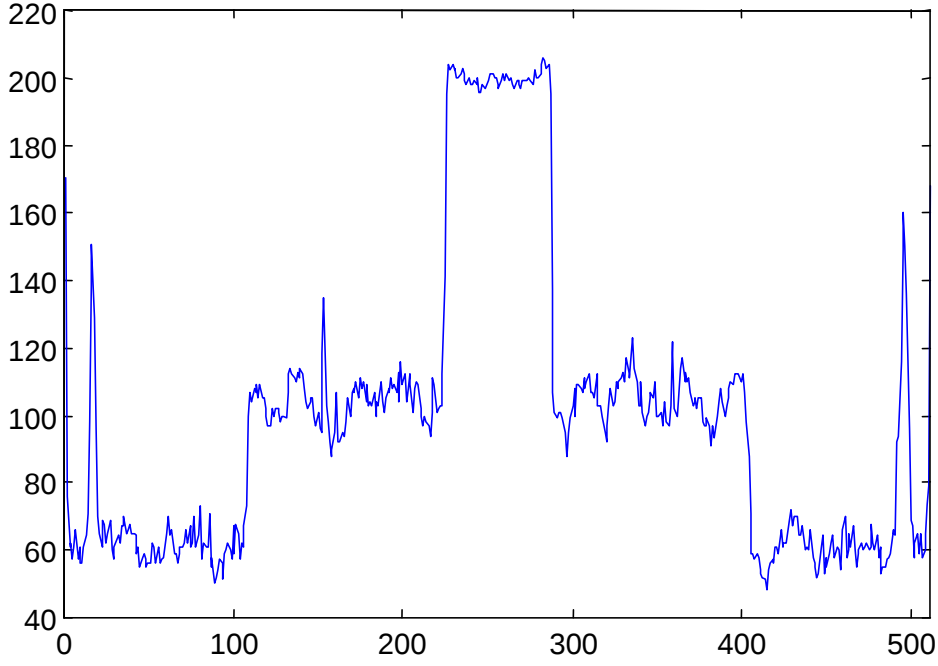


Fig 1.11 Un signal représentant une ligne de l'image de la figure 'Fig 5.2'

Dans le cas bidimensionnel, le passage de l'échelle $j-1$ à l'échelle j se fera selon les quatre équations suivantes :

$$s^j(n, m) = \sum_k \sum_l \tilde{h}\left(\frac{n-k}{2^{j-1}}\right) \tilde{h}\left(\frac{m-l}{2^{j-1}}\right) s^{j-1}(k, l)$$

$$dh^j(n, m) = \sum_k \sum_l \tilde{h}\left(\frac{n-k}{2^{j-1}}\right) \tilde{g}\left(\frac{m-l}{2^{j-1}}\right) s^{j-1}(k, l);$$

$$dv^j(n, m) = \sum_k \sum_l \tilde{g}\left(\frac{n-k}{2^{j-1}}\right) \tilde{h}\left(\frac{m-l}{2^{j-1}}\right) s^{j-1}(k, l);$$

$$dd^j(n, m) = \sum_k \sum_l \tilde{g}\left(\frac{n-k}{2^{j-1}}\right) \tilde{g}\left(\frac{m-l}{2^{j-1}}\right) s^{j-1}(k, l);$$

Nous obtenons, à chaque échelle, une image d'approximation et trois images détails correspondants aux trois directions : horizontale, verticale, diagonale. Chaque image résultat aura les mêmes dimensions que l'image originale (à cause de la non-décimation).

La reconstruction se fait alors selon :

$$\begin{aligned} s^{j-1}(n, m) = & \sum_k \sum_l h\left(\frac{n-k}{2^{j-1}}\right) h\left(\frac{m-l}{2^{j-1}}\right) s^j(k, l) + \sum_k \sum_l h\left(\frac{n-k}{2^{j-1}}\right) g\left(\frac{m-l}{2^{j-1}}\right) dh^j(k, l) \\ & + \sum_k \sum_l g\left(\frac{n-k}{2^{j-1}}\right) h\left(\frac{m-l}{2^{j-1}}\right) dv^j(k, l) + \sum_k \sum_l g\left(\frac{n-k}{2^{j-1}}\right) g\left(\frac{m-l}{2^{j-1}}\right) dd^j(k, l). \end{aligned}$$

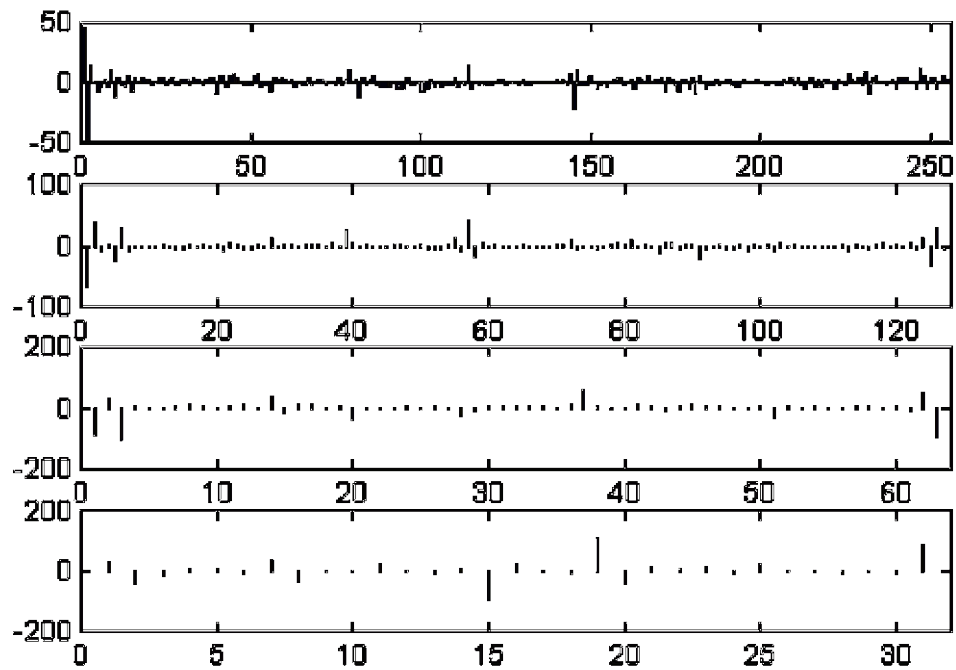


Fig 1.12 Transformation en ondelettes avec décimation du signal précédent , de l'échelle 1, 1^{ère} ligne à l'échelle 4, 4^{ème} ligne

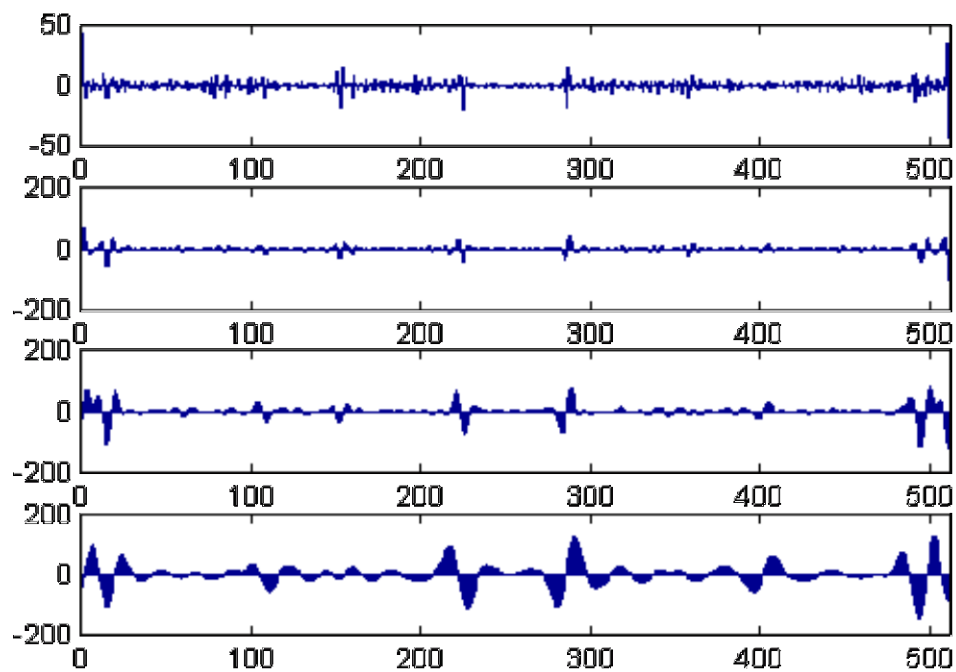


Fig 1.13 Transformation en ondelettes sans décimation du signal précédent , de l'échelle 1, 1^{ère} ligne à l'échelle 4, 4^{ème} ligne

Le signal représenté dans la figure fig1.11 contient des variations subites. Une comparaison entre les représentations en ondelettes obtenues avec décimation fig 1.12 et sans décimation fig1.13 montre que la deuxième permet de traduire les transitions d'un signal dans le domaine des ondelettes d'une manière plus appuyée que la première : ceci grâce à l'évolution progressive des coefficients en ondelettes jusqu'à une valeur maximale locale à l'endroit de la singularité.

Après avoir définie la transformation en ondelettes et les algorithmes qui permettent son obtention, nous allons voir comment vont être utilisés ces coefficients en ondelettes pour l'élimination du bruit dans les images.

2.1 Réduction du bruit par seuillage des coefficients d'ondelette

2.1.1 Définitions

La plupart des méthodes de réduction du bruit se basent sur le modèle additif suivant ; où s est un signal discret de nombre de donnée N et bruité par η .

$$y = s + \eta$$

Le vecteur y représente le signal bruité, le bruit η est un vecteur aléatoire et s est un signal déterministe (tous ces vecteurs ont une dimension N).

On suppose aussi que le bruit à une moyenne nulle $E(\eta) = 0$.

La matrice de covariance de η est définie par $Q = E((\eta - E(\eta))(\eta - E(\eta))^t)$. Sa diagonale sera formée par les variances σ_{ii}^2 . Si cette matrice est diagonale : $E(\eta_i \eta_j) = 0$, si $i \neq j$ alors, le bruit est appelé blanc ou non corrélé (en terme d'auto-corrélation). Si tous les points ont la même densité de probabilité alors il est identiquement distribué $\sigma_i^2 = \sigma^2$.

2.1.2 La transformée en ondelettes d'un signal bruité

La linéarité de la Transformée en Ondelette (T.O) maintient le modèle additif inchangé :

$$w = v + \omega$$

on montre que la matrice covariance S dans le domaine des ondelettes est :

$$S = \tilde{W} Q \tilde{W}^T \quad (2.1)$$

Si la TO est orthogonale et si on a $Q = \sigma^2 I$, alors $S = \sigma^2 I$. (où I est la matrice identité)

Ceci veut dire que la transformée en ondelette orthogonale d'un bruit blanc stationnaire est un bruit blanc stationnaire.

La TO décorrèle les structures d'un signal, et laisse non-corrélé un bruit non-corrélé.

2.1.3 Le seuillage des coefficients en ondelettes:

L'analyse des coefficients en ondelettes d'un signal bruité nous mène à conclure que les petits coefficients sont dominés par le bruit alors que les coefficients avec une grande valeur détiennent plus d'information sur le signal que sur le bruit. En remplaçant les petits coefficients par zéro et en inversant la transformée nous obtenons un signal avec toutes les structures essentielles et moins de bruit. Cette idée est motivée par trois considérations :

- La propriété de la TO de décorrélérer le signal.
- Le bruit affecte tous les coefficients.
- Le niveau du bruit n'est pas élevé.

Le seuillage des coefficients d'ondelette combine simplicité et efficacité. Une première méthode de seuillage, dans laquelle les petits coefficients sont éliminés et les autres sont laissés avec les mêmes valeurs est appelée seuillage dur ou (hard thresholding). Une méthode alternative est le seuillage doux (soft thresholding), les coefficients au-dessus d'un certain seuil sont réduits en valeur absolue; la réduction est égale au seuil.

Bien que le seuillage dur soit une opération qui semble naturelle, la continuité du seuillage doux a des avantages importants et l'analyse de cet algorithme montre des propriétés mathématiques intéressantes. Comme on le verra plus loin certains coefficients bruités peuvent dépasser un seuil, dans le cas du seuillage dur ceci apparaît en sortie comme de fausses structures gênantes. Celles-ci sont réduites avec un seuillage doux.

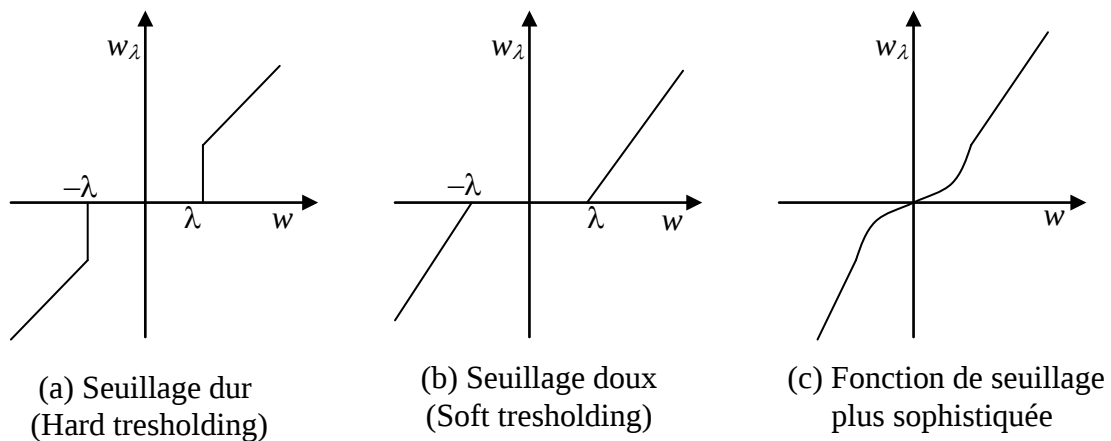


Fig 2.1. Représentation des coefficients d'ondelette après seuillage en fonction des coefficients d'ondelette bruités

Un compromis entre ces deux méthodes qui à la fois préserve les grands coefficients et évite une transition brusque et discontinue a été exploré (firm shrinkage) [Bru97] mais l'inconvénient de ce compromis est la complexité des calculs.

En conclusion le seuillage doux est le meilleur compromis entre temps de calcul et continuité de la transition.

La question centrale qui se pose est comment choisir le seuil qui préserve l'information dans l'image et réalise un débruitage.

Un petit seuil préserve l'image mais n'est pas efficace pour éliminer le bruit, alors qu'un grand seuil produit un résultat excessivement adouci avec une perte d'informations (effet de flou). Beaucoup de littérature est dévolue à ce problème de choix de seuil. Parmi ces méthodes, certaines favorisent l'aspect de douceur comme le seuil universel [Don95b]

[Don94] [Don95a], d'autres visent à minimiser l'erreur quadratique moyenne [Don95c] [Jan97] [Nas96].

2.2 L'aspect non linéaire

L'algorithme de seuillage doit contrôler un compromis entre préservation des structures importantes dans l'image et lissage du bruit indésirable. La dispersion du signal peut être exprimée par une mesure d'entropie, qui doit être minimisée. Cependant l'algorithme assure la préservation de l'information par l'intermédiaire d'un terme traduisant l'erreur de l'énergie (la norme de la différence entre l'entrée et la sortie). Un paramètre de lissage λ permet de contrôler le compromis. L'algorithme doit donc minimiser la quantité suivante :

$$\|w_\lambda - w\|^2 + \lambda \varepsilon(w_\lambda), \quad (2.2)$$

où $\varepsilon(w_\lambda)$ est l'entropie de la sortie. Il existe plusieurs expressions pour celle-ci .

Ne garder que les coefficients qui sont supérieurs à un seuil est une opération non linéaire [Cham98]; elle correspond à la minimisation de l'expression suivante :

$$\sum_{i=1}^N |w_{\lambda i} - w_i|^k + \lambda^k x(w_{\lambda i}),$$

où $x(w_{\lambda i})$ est un label tel que :
$$\begin{cases} x(w_{\lambda i}) = 0 & \text{si } w_{\lambda i} = 0, \\ x(w_{\lambda i}) = 1 & \text{ailleurs,} \end{cases}$$

ceci correspond au seuillage dur

$$\begin{cases} w_{\lambda i} = w_i & \text{si } |w_i| \geq \lambda, \\ w_{\lambda i} = 0 & |w_i| < \lambda, \end{cases}$$

Pour $k=2$, l'expression se ramène à (2.2), où l'entropie serait le nombre de coefficients différents de zéro. $\varepsilon(w_\lambda) = N_1 = \text{card}\{i = 1, \dots, N / w_{\lambda i} \neq 0\}$,

Le seuillage doux s'ensuit si on utilise la norme ℓ_1 comme mesure de dispersion. On minimise dans ce cas la quantité suivante :

$$\sum_{i=1}^N ((w_{\lambda i} - w_i)^2 + 2\lambda |w_{\lambda i}|),$$

ceci nous indique que l'entropie est égale à :

$$\varepsilon(w_\lambda) = 2 \sum_{i=1}^N |w_{\lambda i}|,$$

2.3 Erreur quadratique moyenne (EQM) et fonction risque

2.3.1 Définition :

Le seuillage des coefficients en ondelettes peut être vu comme une opération de lissage du signal original. L'objectif dans l'approximation de l'information débruitée est de contrôler le compromis entre la préservation des structures et l'élimination du bruit indésirable par lissage. Le meilleur choix minimise l'erreur entre le résultat et la fonction sans bruit (en réalité fonction inconnue).

Si y_λ est le résultat d'un seuillage avec un seuil λ et s est le vecteur de données sans bruit le bruit résiduel dans y_λ est: $\eta_\lambda = y_\lambda - s$,

L'erreur quadratique moyenne (EQM) est définie alors comme :

$$R(\lambda) = \text{EQM}(\lambda) = \frac{1}{N} \|\eta_\lambda\|^2.$$

Comme il apparaît de cette expression, l'EQM $R(\lambda)$, est une fonction de seuil λ ; c'est aussi une variable aléatoire car elle dépend du bruit. La valeur moyenne de cette erreur est appelée fonction risque.

Etant donné que l'EQM utilise dans sa définition la fonction sans bruit inconnue s , elle ne peut être calculée exactement. On a recours en pratique à une estimation de cette quantité.

Une transformation orthogonale préserve la norme ℓ_2 pour cela , on peut écrire :

$$R(\lambda) = \frac{1}{N} \|\omega_\lambda\|^2 \quad \text{où} \quad \omega_\lambda = w_\lambda - v = \tilde{W}(y_\lambda - s)$$

On peut donc effectuer toutes les opérations dans le domaine des ondelettes. Si la transformation est biorthogonale, il n'y a pas d'équivalence avec le domaine d'origine (le domaine spatial). Cependant les calculs et la minimisation en fonction des coefficients d'ondelette donnent des résultats satisfaisants.

En outre, étant donné qu'une image malgré qu'elle soit mise sous forme de matrice de pixels n'est pas regardée comme telle par l'œil humain mais plutôt regardée dans une base multi-échelle, il serait plus judicieux de minimiser l'EQM dans le domaine des ondelettes. Ceci constituerait une meilleure traduction du critère de qualité de la perception humaine et nous évite d'inverser la transformée à chaque fois que nous voulons évaluer le résultat du traitement. La transformation inverse sera utilisée seulement pour obtenir le résultat final de l'algorithme.

2.3.2 Variance et biais :

Les coefficients en ondelettes de l'entrée sont des estimations non biaisées des coefficients en ondelettes de la fonction sans bruit, mais la variance de cette estimation est plutôt élevée. Remplacer les plus petits coefficients par zéro réduit cette variance mais au prix d'un biais plus important.

$$\text{biais}^2(\lambda) = \|E(w_\lambda) - v\|^2.$$

$$\text{variance}(\lambda) = E\|w_\lambda - E(w_\lambda)\|^2.$$

ce qui donne :

$$E(R(\lambda)) = \frac{1}{N} \left(\|E(w_\lambda) - v\|^2 + E\|w_\lambda - E(w_\lambda)\|^2 \right)$$

$$\text{Risque} = \text{biais}^2 + \text{variance}.$$

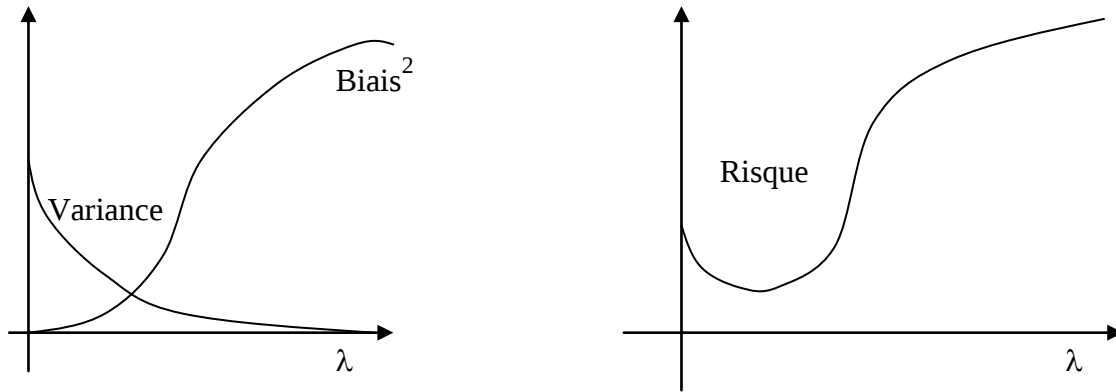


Fig 2.3 Biais , variance et risque en fonction du seuil

Les méthodes qui visent à enlever tout le bruit d'une manière catégorique risquent d'éliminer aussi une partie importante du signal.

$$\lim_{\lambda \rightarrow \infty} \text{variance}(\lambda) = 0$$

Si tous les coefficients sont éliminés, la variance sera réduite à zéro : tout le bruit est supprimé, mais le signal aussi. Dans ce cas, le biais est égal à la racine de l'énergie du signal sans bruit ;

$$\lim_{\lambda \rightarrow \infty} \text{biais}^2(\lambda) = \|v\|^2$$

Après avoir défini l'EQM, l'objectif est de trouver un compromis entre le lissage et la préservation de structures en cherchant un compromis à l'aide de la fonction risque entre variance et biais (tous deux sont inconnus). Une solution avec variance minimale est probablement régulière et une bonne préservation de contours s'accompagne d'un faible

biais ; d'où le meilleur compromis entre variance et biais est la valeur du seuil pour laquelle on a le minimum de risque.

2.3.3 Contribution à la fonction risque de chaque coefficient :

Une étude [Jan 2000] a montré que le choix d'un seuil élevé est souhaité si on considère seulement les coefficients qui s'approchent de zéro, et d'un autre coté un seuil avec une petite valeur est souhaité si on considère seulement les coefficients élevés. Finalement le seuil optimal dépend de la distribution des coefficients.

La contribution à la fonction risque de chaque coefficient est :

$$r(v, \lambda) = [2(\sigma^2 + \lambda^2) - v^2] + [\Phi(\lambda - v) + \Phi(\lambda + v)](v^2 - \sigma^2 - \lambda^2) - \sigma^2[(\lambda - v)\phi(\lambda + v) + (\lambda + v)\phi(\lambda - v)]$$

où : λ est le seuil, v est le coefficient d'ondelette sans bruit, σ le niveau du bruit, ϕ distribution du bruit (loi de Gauss) et Φ fonction de répartition du bruit.

La représentation de cette contribution en fonction du seuil (Fig 2.3) pour différentes valeurs de v montre que les coefficients avec peu d'information préfèrent un seuil élevé contrairement aux larges coefficients qui favorisent le choix d'un petit seuil.

Afin de trouver les minima du risque EQM , on calcule la dérivée :

$$\frac{\partial r}{\partial \lambda}(v, \lambda) = 2\lambda[2 - \Phi(v + \lambda) - \Phi(-v + \lambda)] - 2\sigma^2[\phi(v + \lambda) + \phi(-v + \lambda)], \quad (2.3)$$

Un cas important à souligner se présente lorsque $v=0$, dans ce cas la dérivée serait toujours négative et s'approcherait de zéro quand $\lambda \rightarrow \infty$. Ceci signifie que le seuil optimal pour les coefficients nuls est égal à l'infini.

Pour avoir une idée plus générale, on calcule sa dérivée par rapport au coefficient v :

$$\frac{\partial}{\partial v} \left(\frac{\partial r}{\partial \lambda}(v, \lambda) \right) = 2v[\phi(v + \lambda) + \phi(-v + \lambda)] = \begin{cases} \geq 0 & \text{si } v \geq 0, \\ \leq 0 & \text{si } v \leq 0, \end{cases}$$

En conséquence, pour un seuil donné, $\frac{\partial r}{\partial \lambda}$ a un minimum pour $v=0$.

2.4 Propriété asymptotique du seuil à minimum de risque :

2.4.1 Motivation

Dans le prochain chapitre, il est montré que dans le sens du minimum de risque, la méthode de validation croisée généralisée GCV fournit asymptotiquement le seuil optimal. Cette propriété est motivante pour l'utilisation de la GCV dans l'évaluation du seuil. Pour prouver

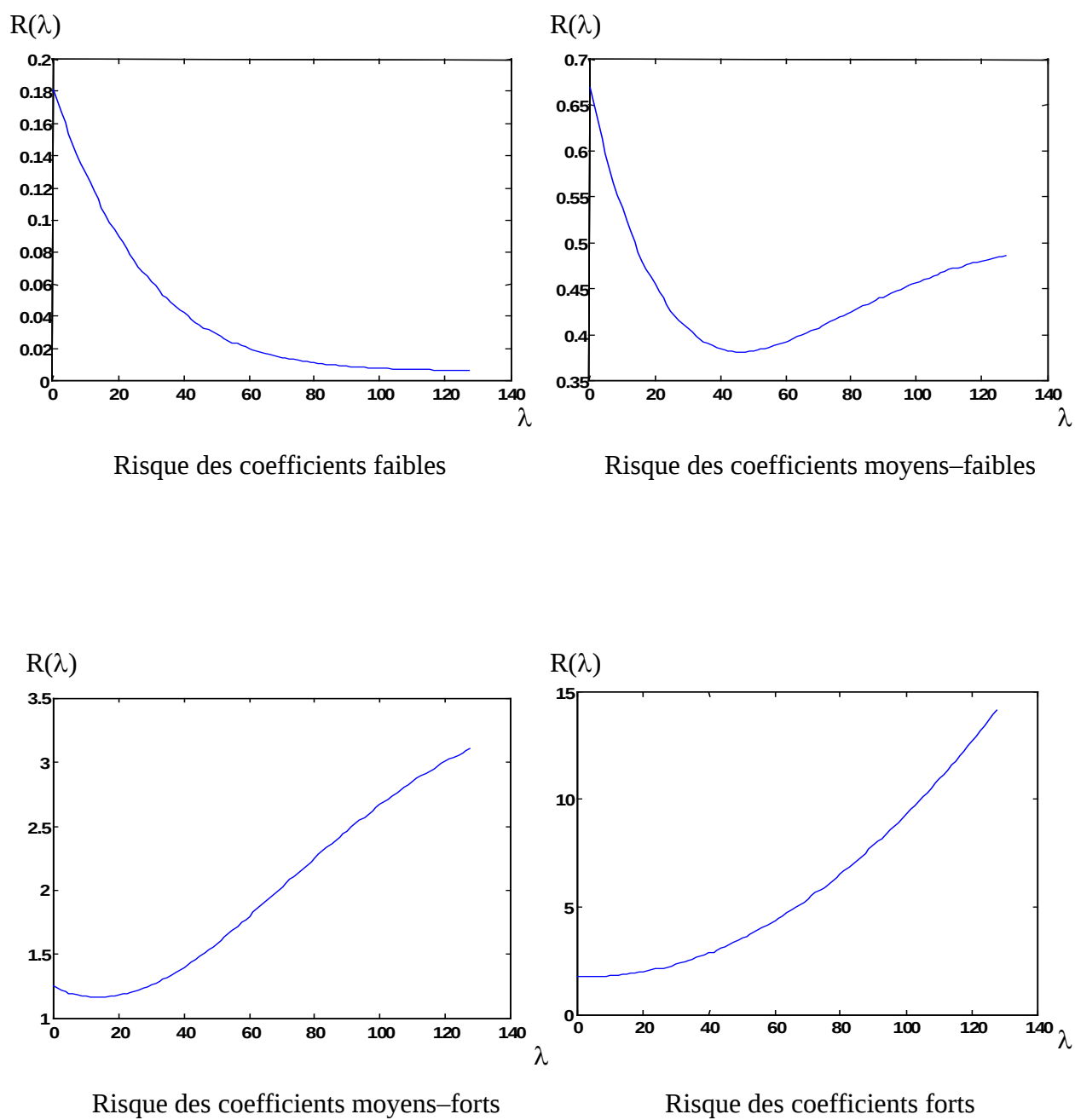


Fig 2.3 Contribution à la Fonction risque en fonction du seuil pour différentes valeurs de coefficients d'ondelette

cette optimalité asymptotique, on doit tout d'abord connaître quel est le comportement du seuil optimal quand le nombre de données tend vers l'infini.

2.4.2 Equivalence asymptotique :

a) Définition :

Deux fonctions $f(x)$ et $g(x)$ sont dites asymptotiquement équivalentes (on notera $f(x) \sim g(x)$) quand x tend vers l'infini, si et seulement si :

$$\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = 1$$

b) Propriétés

Si x tend vers l'infini, et si $f(x) \sim g(x)$, on a :

1. $f(x)h(x) \sim g(x)h(x)$.
2. Si $f(x) \sim h(x)$, alors $h(x) \sim g(x)$.
3. Si $h(x) = o(f(x))$, alors $f(x) \pm h(x) \sim g(x)$.
4. Si $\lim_{x \rightarrow \infty} f(x) \neq 1 \neq \lim_{x \rightarrow \infty} g(x)$, et si les deux fonctions sont différentiables
 $\log(f(x)) \sim \log(g(x))$.

2.4.3 Comportement asymptotique pour les fonctions polynomiales par parties :

Pour le cas de fonctions polynomiales par parties, on suppose que l'ondelette analysante a un nombre de moments nuls plus grand que le plus haut degré des polynômes ; en conséquence les coefficients en ondelettes seront nuls s'ils ne coïncident pas avec des singularités. On suppose que le nombre de singularités est fini sur l'intervalle $[0,1]$.

Théorème : [Don94]

Supposons que $s(x)$ est une fonction polynomiale par partie sur $[0,1]$, $v_{i,k}$ ($i=1 \dots N$) sont les coefficients en ondelettes de s , on note $w_i = v_i + \omega_i$ les coefficients en ondelette bruités et $R(\lambda)$ la fonction d'EQM. Si λ^* minimise $E(R(\lambda))$ alors pour $N \rightarrow \infty$:

$$\lambda^* \sim \sqrt{2 \log N \sigma}$$

Preuve

Soit :

$w = v + \omega$ les coefficients en ondelette du signal s ,

On note :

$I_0 = \{\text{ensemble des coefficients } v \text{ nuls}\}$

$I_1 = \{\text{ensemble des coefficients } v \text{ non-nuls}\}$

$M_0 = \text{nombre d'éléments de } I_0.$

$M_1 = \text{nombre d'éléments de } I_1.$

Comme précédemment expliqué, on peut écrire :

$$R(\lambda) = \sum_{i=1}^N r(v_i, \lambda),$$

d'où

$$R'(\lambda) = M_0 \frac{\partial r}{\partial \lambda}(0, \lambda) + \sum_{i \in I_1} \frac{\partial r}{\partial \lambda}(v_i, \lambda),$$

si on divise l'ensemble I_1 en deux sous ensembles :

- I_1' dans lequel se trouve les coefficients v_i non-nuls tels que : $\frac{\partial r}{\partial \lambda}(v_i, \lambda)$ est négatif.
- I_1'' est l'ensemble complémentaire de I_1' dans I_1 .

M_1' et M_1'' sont respectivement les nombres d'éléments des sous-ensembles I_1' I_1'' .

Pour le minimum de $R(\lambda)$, on a $R'(\lambda)=0$, d'où :

$$-M_0 \frac{\partial r}{\partial \lambda}(0, \lambda^*) + \sum_{i \in I_1'} \frac{\partial r}{\partial \lambda}(v_i, \lambda^*) = \sum_{i \in I_1''} \frac{\partial r}{\partial \lambda}(v_i, \lambda^*),$$

le coté droit de l'équation est équivalent à $M_1'' \frac{\partial r}{\partial \lambda}(v_i, \lambda^*)$. Ceci est déduit de l'équation 2.3 quand

$v \rightarrow \infty$. Dans le coté gauche on néglige $\sum_{i \in I_1'} \frac{\partial r}{\partial \lambda}(v_i, \lambda^*)$, car $\frac{M_1'}{M_0} \sim \frac{\log N}{N} \rightarrow 0$;

En conclusion nous avons

$$\frac{M_0}{M_1''} \sim \frac{\lambda^*}{-\frac{\partial r}{\partial \lambda}(0, \lambda^*)}$$

à partir de l'équation 2.3 on peut trouver que

$$\frac{\partial r}{\partial \lambda}(0, \lambda^*) \sim -4\sigma^4 \frac{\phi(\lambda^*)}{\lambda^{*2}}.$$

d'où

$$\frac{M_0}{M_1''} \sim \frac{2\lambda^{*3}}{4\sigma^4\phi(\lambda^*)}$$

$$\frac{\sigma^3 M_0}{M_1''} \sim \sqrt{\frac{\pi}{2}} e^{\lambda^{*2}/2\sigma^2} \lambda^{*3},$$

$$3\log \sigma + \log M_0 - \log M_1'' \sim \frac{\lambda^{*2}}{2\sigma^2} + 3\log \lambda^* + \frac{1}{2}\log \frac{\pi}{2},$$

En ne retenant que les termes essentiels , l'expression se réduit de la manière suivante :

$$\log M_0 \sim \frac{\lambda^{*2}}{2\sigma^2}$$

$$\log(N - M_1) \sim \frac{\lambda^{*2}}{2\sigma^2}$$

$$2\sigma^2 \log N \sim \lambda^{*2}$$

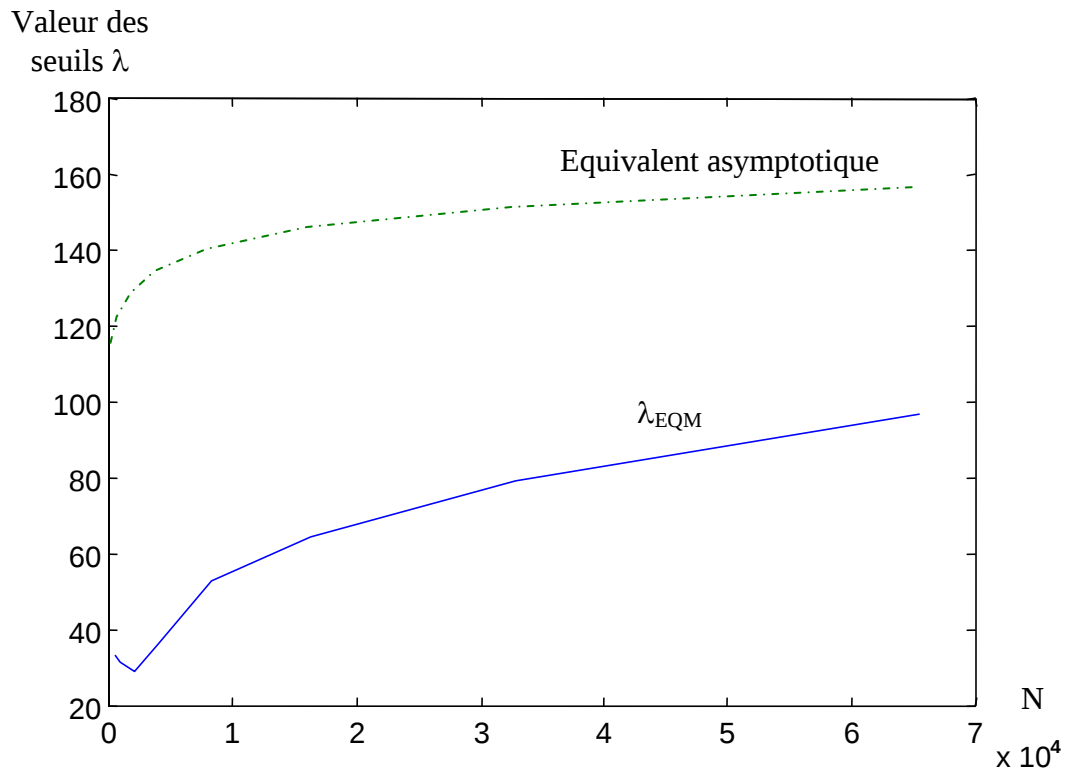


Fig 2.4. Seuil optimal (minimum EQM) (trait continu), et son équivalent asymptotique $\sigma\sqrt{2\log(N)}$ (trait discontinu), en fonction du nombre de données N

2.5 Seuil universel :

Il est sûr que l'étude asymptotique du seuil à minimum de risque ne révèle pas tout sur la valeur du seuil optimal. Cette valeur dépend de tous les coefficients tandis que la valeur asymptotique ne dépend que de N et σ (N nombre de donnée et σ le niveau du bruit).

$$\lambda_{UNIV} = \sqrt{2\log(N)}\sigma.$$

Dans l'étude asymptotique précédente, il n'a jamais été question d'utiliser cette valeur asymptotique comme seuil. Cependant cette valeur est bien connue dans les travaux relatifs au seuillage par la T.O et elle y est utilisée comme seuil et non pas seulement comme équivalent asymptotique elle est connue sous le nom de seuil universel. Cette appellation reflète l'idée que ce seuil est valide pour tous les signaux de longueur N pourvu que ces signaux soient suffisamment réguliers : c'est une valeur générale du seuil. Donoho et Johnstone [Don94] ont démontré beaucoup de propriétés optimales pour ce choix, nous les citerons brièvement :

a) Imitation du seuillage optimal :

Si on suppose qu'on peut deviner les coefficients non bruités en les choisissant au-dessus du seuil optimal et en éliminant les autres, la sélection optimale de coefficients en ondelette (symbolisée SO dans les inégalités suivantes) réaliserait un résultat meilleur que les méthodes polynomiales par partie (PP):

$$\text{Risque}(\text{SO}) \leq (A \log N + B)(\text{Risque}(\text{PP}))$$

où : A, B sont des constantes réelles.

D'un autre côté, le seuillage en utilisant le seuil universel donne des résultats aussi bons que la reconstruction sélective optimale.

$$\text{Risque}(\lambda_{UNIV}) \leq (2 \log N + 1)(\text{Risque}(\text{SO}) + \frac{\sigma^2}{N}).$$

On déduit que la procédure de seuillage par le seuil universel, en s'appuyant seulement sur les données, atteint, à un facteur $\log^2 N$, les performances des méthodes polynomiales équipées d'un oracle.

b) La propriété Minimax :

Soit λ_{mM} le seuil minimax (le plus grand seuil parmi ceux qui minimisent le risque relatif maximum par rapport au risque de la sélection optimale).

λ_{mM} est asymptotiquement équivalent à $\sqrt{2 \log N}$

c) Adaptivité et optimalité

Les résultats précédents relient les performances du seuil universel à la sélection idéale des coefficients : ceci exprime le fait que si la sélection optimale donne de bons résultats alors le seuillage doux par un seuil universel donne des résultats qui s'en approchent. Ceci nous amène à se demander quand est-ce que la reconstruction sélective est une bonne méthode pour la réduction du bruit. Nous devons nous attendre à de bons résultats quand la représentation en ondelette est éparse. C'est justement le cas pour les signaux doux (réguliers) par parties. Pour caractériser cette douceur, on ne peut utiliser l'espace de Hölder car une singularité détruit la douceur, on utilise plutôt l'espace de Besov $B_{p,q}^s$ où les paramètres p , q et s mesurent différents aspects de la douceur.

Si une fonction F s'étend dans un tel espace le risque du seuil universel atteint, à un facteur logarithmique près, le risque du seuil minimax.

$$\sup_{f \in F} \text{Risque}(\lambda_{UNIV}) \leq C \log N \cdot \inf_{\hat{f}} (\sup_{f \in F} (\text{Risque}(\hat{f})))$$

Le seuil universel a la propriété d'aboutir à ces résultats sans connaissances a priori sur la douceur du signal, cependant l'estimation optimale dépend clairement d'une mesure de cette douceur. C'est pourquoi la procédure de seuillage est dite adaptative à une douceur inconnue [Don95c].

d) La douceur

L'espace de Besov est muni d'une norme pour mesurer la douceur d'une fonction, tout en étant flexible pour mesurer les singularités : une singularité isolée ne pose pas de problème pour la douceur au sens de Besov.

Il a été prouvé que la reconstruction, en utilisant le seuil universel, produit avec une grande probabilité, un résultat aussi doux que le signal sans bruit [Don95b]. Cette garantie de douceur n'existe pas pour les seuillages se basant sur une approche de minimisation de risque. Les résultats de ces algorithmes contiennent des structures bruitées « blips ». Une amélioration de ce seuillage sera présentée pour remédier à ce défaut.

Nous avons présenté dans ce qui a précédé une méthode d'élimination du bruit agissant dans le domaine des ondelettes et se basant sur le seuillage. Nous avons opté pour le seuil qui minimise l'EQM. Nous avons montré que ce seuil est asymptotiquement équivalent au seuil universel. Le chapitre suivant exposera les méthodes SURE et GCV qui permettront d'estimer l'EQM et le seuil à EQM minimale.

Dans chapitre précédent, on a étudié les propriétés du seuillage à risque minimal. En pratique, l'EQM ne peut jamais être évaluée exactement parce qu'elle dépend des coefficients non bruités dans le calcul. C'est pourquoi, nous avons besoin d'estimer cette fonction.

3.1 Estimateur SURE de l'EQM

On veut obtenir une estimation de l'EQM (qui est basée sur des variables inconnues). Pour cela on considère l'effet du seuillage sur les données en entrée. On définit la fonction:

$$F(\lambda) = \frac{\sum_{i=1}^N (w_{\lambda_i} - w_i)^2}{N} = \frac{1}{N} \|w_{\lambda} - w\|^2.$$

La valeur moyenne de cette fonction est :

$$E(F(\lambda)) = E\left(\frac{1}{N} (\|w - v\|^2 + \|v - w\|^2 + 2 \langle w - v, v - w_{\lambda} \rangle)\right).$$

$$E(F(\lambda)) = \sigma^2 + E(R(\lambda)) - \frac{2}{N} E(w, w_{\lambda}).$$

Le troisième terme peut s'exprimer d'une autre façon grâce au lemme suivant:

Lemme [Ste81]

Si la densité $\phi(w_i)$ est gaussienne, pour le seuillage doux on a :

$$E(w_i, w_{\lambda_i}) = \sigma^2 P(|w_i| > \lambda)$$

On définit la fonction suivante aussi :

$$\mu_1(\lambda) = \frac{\sum_{i=1}^N E(w_i, w_{\lambda_i})}{N\sigma^2}.$$

A partir du lemme on peut écrire :

$$\mu_1(\lambda) = \frac{\sum P(|w_i| > \lambda)}{N}.$$

On introduit également la matrice :

$$D_{ij} = \frac{\partial w_{\lambda_i}}{\partial w_j}$$

On remarque que si $i \neq j$ on a $D_{ij} = 0$, et pour $i = j$ on a $D_{ii} = \begin{cases} 0 & \text{si } |w_i| < \lambda. \\ 1 & \text{ailleurs.} \end{cases}$

Pour cela, la trace de D , soit $\text{Tr}(D)$, serait égale au nombre de données au dessus du seuil, soit :

$$\text{Tr}(D) = \text{card}\{i / w_{\lambda i} \neq 0\}$$

$$\text{Tr}(D) = N - N_0.$$

où N_0 est égal au nombre de données au dessous du seuil $N_0 = \text{card}\{i / w_{\lambda i} = 0\}$.

Si A est une matrice telle que

$$A_{ij} = \frac{\partial y_{\lambda i}}{\partial y_i}$$

Il est facile de voir que $A = \tilde{W}^{-1} D \tilde{W}$. où $\tilde{W}$ est la matrice T.O, si la transformation en ondelette est orthogonale alors on peut écrire :

$$A = \tilde{W}^{-1} D (\tilde{W}^{-1})^T$$

Comme $\tilde{W}$ n'est pas singulière, on a : $\text{Tr}(A) = \text{Tr}(D)$.

Avec cette notation et puisque pour les variables Bernoulli on a $E(D_{ii}) = P(D_{ii} = 1)$, on peut écrire comme suit :

$$\mu_1 = \frac{1}{N} \sum P(D_{ii} = 1) = \frac{1}{N} \sum E(D_{ii}) = \frac{\text{Tr}(E(A))}{N}$$

A partir de $\langle w, w_{\lambda} \rangle$ (incalculable en pratique), on a abouti à $\sigma^2 \text{Tr}(A)$.

On peut écrire comme approximation de $R(\lambda)$:

$$\text{SURE}(\lambda) = F(\lambda) - \sigma^2 + \frac{2\sigma^2}{N} \text{Tr}(A).$$

Cette estimation a été découverte au départ par Stein [Ste81] et porte son nom SURE pour Stein's Unbiased Risk Estimator. A partir de l'appellation on devine que l'estimateur est non biaisé, cette propriété n'est pas asymptotique comme dans le cas de la GCV ; le nombre de coefficients ne joue pas de rôle dans le calcul de la valeur moyenne. L'estimateur peut être calculé dans le domaine d'origine, mais comme :

$$\text{Tr}(A) = \text{Tr}(D) = N - N_0.$$

Il est plus facile de compter le nombre de coefficients nuls dans le domaine des ondelettes.

3.2 SURE Adaptive

Contrairement au seuil universel, le seuil SURE dépend directement du signal à traiter. On doit donc s'attendre à un seuil plus adaptatif. Donoho et Johnston [Don95c] ont expliqué

comment que la procédure SURE s'adapte automatiquement à la classe de douceur (espace de Besov) à laquelle probablement les données non bruitées appartiennent.

La méthode atteint asymptotiquement le comportement minimax à une constante près ; c'est à dire que si une fonction non bruitée appartient à un espace de Besov $F = B_{p,q}^s$ l'algorithme SURE agit comme une procédure minimax : il s'approche de l'estimateur qui minimise le pire cas de risque dans l'espace de fonction.

$$\sup_{f \in F} \text{Risque}(\text{SURE}) \leq C \inf_{\hat{f}} \sup_{f \in F} \text{Risque}(\hat{f}).$$

C dépend des paramètres de douceur p , q et s de l'espace de Besov .

Voici pourquoi la méthode s'adapte automatiquement au type de douceur de la fonction non bruitée. La procédure de seuillage ne nécessite pas des connaissances supplémentaires sur la douceur du signal pour trouver un seuil qui s'approche du meilleur seuil.

Cette performance presque minimax de SURE est meilleure que celle du seuil universel d'un facteur logarithmique.

3.3 Validation croisée ordinaire

Cette partie introduit le principe de la validation croisée (cross validation)[Wah90]. Notre but est de minimiser la fonction d'erreur basée sur le signal exact inconnu. Pour cela nous cherchons un compromis entre lissage et préservation de l'information. Nous supposons que le signal original est régulier jusqu'à une certaine mesure , ce qui signifie qu'une valeur f_i peut être approximée par une combinaison linéaire des valeurs qui lui sont voisines. En considérant une combinaison linéaire de y_i , qui ne dépend pas de y_i , on peut éliminer le bruit dans cette partie. Du moment qu'on remplace un point par une moyenne pondérée de ces voisins , le bruit en ce point est lissé et on aboutit comme ça à un résultat relativement moins bruité. Cette valeur peut être utilisée pour calculer une approximation de la fonction risque $R(\lambda)$.

Pour étudier la préservation des structures, on calcule le résultat du seuillage Λ du signal modifié , dans lequel la $i^{\text{ème}}$ composante y_i est remplacé par $\tilde{y}_i$.

$$\text{ie : } \tilde{y} = \Lambda([y_1, y_2, \dots, y_{i-1}, \tilde{y}_i, y_{i+1}, \dots, y_N]^T).$$

On considère la capacité de $\tilde{y}_{\lambda_i}$ à prédire la valeur de y_i comme une mesure d'optimalité du choix du seuil. Pour de faibles valeurs du seuil λ , la différence $y_i - \tilde{y}_{\lambda_i}$ est dominée par le

bruit, alors que pour de larges valeurs du seuil le signal est déformé. On répète la procédure et on calcule, pour exprimer le compromis, la grandeur :

$$OCV = \frac{1}{N} \sum_{i=1}^N (y_i - \tilde{y}_{\lambda_i})^2.$$

Cette fonction est appelée « validation croisée ordinaire » (ordinary cross validation), cette appellation nous indique que, pour calculer un point, on utilise les valeurs d'autres points (les voisins). L'OCV nécessite N procédures de seuillage y compris la TO et la TO inverse.

Plusieurs formules de combinaisons sont possibles pour $\tilde{y}_i$, la plus évidente est :

$$\tilde{y}_i = \frac{1}{2}(y_{i-1} + y_{i+1}).$$

Pour ce dernier choix de $\tilde{y}_i$ on peut écrire :

$$y_i - \tilde{y}_{\lambda_i} = \frac{y_i - y_{\lambda_i}}{1 - a_i^*}, \text{ avec : } a_i^* = \frac{y_{\lambda_i} - \tilde{y}_{\lambda_i}}{y_i - \tilde{y}_{\lambda_i}} = \frac{y_{\lambda_i} - \tilde{y}_{\lambda_i}}{y_i - \tilde{y}_i} \approx \frac{\partial y_{\lambda_i}}{\partial y_i} = A_{ii}.$$

Donc on a

$$OCV \approx \frac{1}{N} \sum_{i=1}^N \alpha_i^2(\lambda) (y_i - y_{\lambda_i})^2, \text{ avec } \alpha_i(\lambda) = \frac{1}{1 - A_{ii}}$$

Cette expression ne peut être évaluée dans le domaine des ondelettes et le calcul de A en utilisant $A = \tilde{W}^{-1} D \tilde{W}$ est coûteux puisque pour chaque évaluation, on a besoin d'une TO inverse. En plus, les expérimentations ont indiqué que l'évaluation est un problème mal posé, spécialement pour les petites valeurs du seuil, quand la plupart des A_{ii} sont égaux à 1. Pour accélérer la procédure, on peut utiliser une sorte de valeur moyenne de $\alpha_i(\lambda)$:

$$\alpha_i(\lambda) = \alpha(\lambda) = \frac{1}{\frac{1}{N} \sum_{i=1}^N 1 - A_{ii}} = \frac{1}{\frac{1}{N} (Tr(I - A))} = \frac{1}{\frac{1}{N} (Tr(I - D))}.$$

Ceci nous donne la fonction dite de la validation croisée généralisée « GCV » qui peut être évaluée et minimisée dans le domaine des ondelettes.

3.4 Validation Croisée Généralisée GCV (Generalized cross validation)

La GCV [Wey98][Nas96][Jan97] est une fonction du seuil définie par :

$$GCV(\lambda) = \frac{\frac{1}{N} \|y - y_{\lambda}\|^2}{\left[\frac{Tr(I - A)}{N} \right]^2} = \frac{F(\lambda)}{\left[\frac{N_0(\lambda)}{N} \right]^2}.$$

Cette fonction peut atteindre une valeur infinie pour $N_0=0$. Ceci se produit difficilement pour des signaux affectés par un bruit gaussien si $\lambda > 0$ mais on ne peut pas affirmer que $P(N_0=0)=0$ si $N_0 = 0$. Pour éviter tout problème, on pose $GCV=0$ si $N_0=0$.

Si la transformation en ondelettes est orthogonale la même formule peut être utilisée dans le domaine des ondelettes

$$GCV(\lambda) = \frac{\frac{1}{N} \|w - w_\lambda\|^2}{\left[\frac{N_0(\lambda)}{N} \right]^2}.$$

La minimisation de cette fonction peut être faite dans le domaine des ondelettes. Le dénominateur peut être facilement évalué en effectuant le comptage des coefficients dont la valeur est au dessus du seuil.

3.5 Comportement asymptotique

On démontre [Wey98][Jan2000] que si $\lambda^* = \arg \min R(\lambda)$ et $\tilde{\lambda} = \arg \min GCV(\lambda)$ alors pour $N \rightarrow \infty$ les deux procédures de minimisation donnent des résultats de même qualité (équivalents), soit :

$$\frac{E(R(\tilde{\lambda}))}{E(R(\lambda^*))} \rightarrow 1$$

et au voisinage de λ^* $E(GCV(\lambda))$ serait une translation verticale de $E(R(\lambda))$ avec une valeur de σ^2 .

$$E(R(\lambda)) + \sigma^2 \approx E(GCV(\lambda));$$

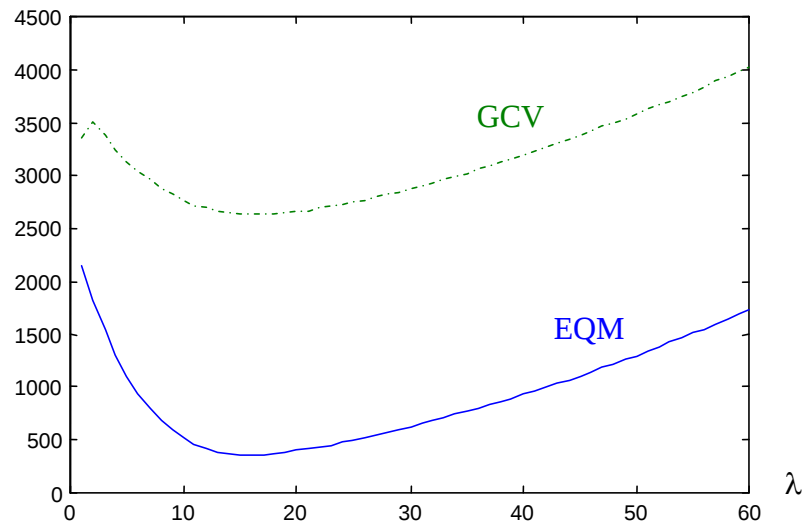


Fig 3.1 GCV et EQM de la fonction résultante en fonction du seuil

3.6 GCV pour un nombre de données fini :

L'optimalité asymptotique assure un bon résultat dans la plupart des cas. Il suffit seulement que le nombre de données N utilisées pour estimer la GCV soit assez élevé. Pour les applications d'élimination de bruit dans les signaux, nous avons conclu, d'après nos expérimentations qu'un nombre minimum de données égal à $N = 1000$, est requis pour que la procédure de minimisation puisse estimer correctement le seuil optimal.

Nous illustrons dans les graphes Fig 3.2 et Fig 3.3 suivants l'effet de la réduction du nombre de données sur la capacité de la GCV à estimer l'EQM. Nous constatons la détérioration des qualités de la GCV dans l'estimation de l'EQM pour un faible nombre de données. Néanmoins le minimum de GCV peut être accepté comme une estimation du minimum de l'EQM. Mais il est clair que même cette possibilité est perdue quand le nombre de données se réduit encore plus.

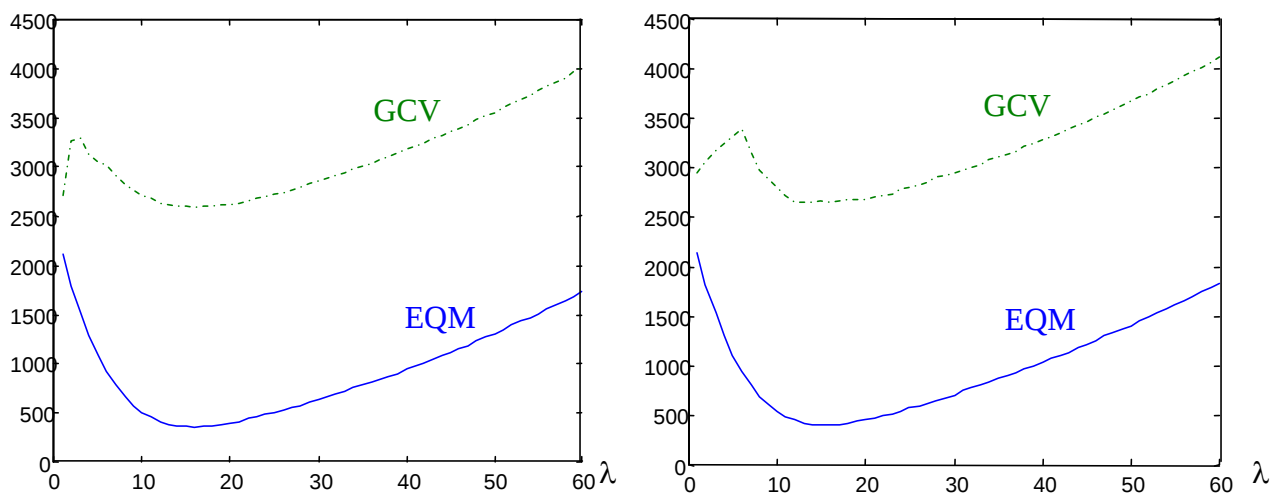


Fig 3.2 GCV et EQM de la fonction résultante en fonction du seuil pour $N=16384$ (gauche) et $N=4096$ (droite)

3.7 Procédure de minimisation :

La minimisation se fait dans le domaine des ondelettes et en utilisant la procédure de minimisation de Fibonacci [Fle81]; il n'est pas nécessaire de calculer le minimum avec précision, une approximation avec une erreur de 10^{-4} suffit du fait que la courbe de EQM en

fonction de λ est régulière dans le voisinage du minimum. Cette procédure atteint une approximation avec la tolérance voulue après 20 itérations en moyenne.

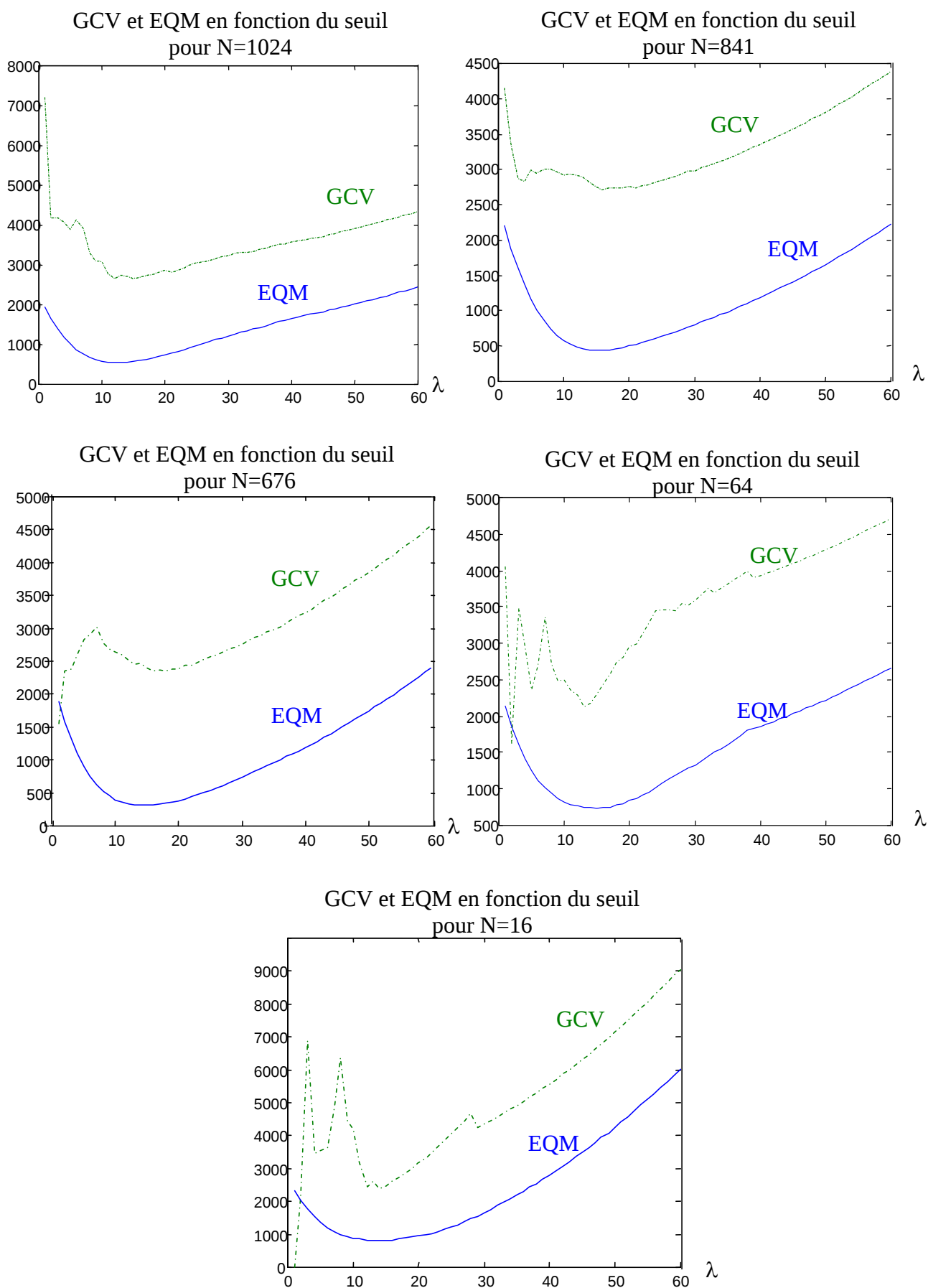


Fig 3.3 Représentation de l'EQM (trait continu) et de son estimation GCV (trait

Nous avons présenté deux méthodes d'estimation de l'EQM (SURE et GCV) qui peuvent s'exprimer dans le domaine des ondelettes afin de calculer le seuil à EQM minimale. On a vu que le seuillage en utilisant ce seuil permet d'aboutir à un résultat plus adaptatif à la douceur du signal. On a apprécié la capacité de la méthode GCV à estimer l'EQM sans connaissance a priori sur le niveau du bruit à condition que celui-ci soit stationnaire et non-corrélé, et que la transformation en ondelette soit orthogonale. Dans le prochain chapitre une extension de cette méthode au cas d'un bruit coloré et d'une transformation en ondelettes non-orthogonale sera proposée.

Nous avons exploité le caractère éparé d'une représentation en ondelette pour justifier le seuillage comme méthode de filtrage et la GCV comme méthode d'évaluation d'un seuil asymptotique optimal. Jusqu'à présent, nous n'avons pas utilisé une autre caractéristique, qui est une voie naturelle de support des ondelettes « l'idée de la multirésolution ». Une représentation en ondelette n'est pas seulement dispersée mais aussi multi-échelle, la combinaison de ces deux approches crée plus de possibilités car :

Premièrement, un signal dans la plupart du temps a des caractéristiques différentes dans chaque échelle ; on considère l'échelle comme une grandeur inversement proportionnelle à la fréquence. On peut utiliser le caractère multi-échelle de cette représentation pour rendre plus adaptatives certaines opérations de traitement du signal et de l'image ; pas seulement en traitant chaque niveau indépendamment mais en traitant un niveau en considérant les informations dans les niveaux adjacents[Cro98][Xu94].

Deuxièmement, si le bruit est coloré au lieu d'être blanc, la répartition du bruit dépend du niveau, d'où un seul seuil ne va pas éliminer le bruit correctement. Un seuillage avec des valeurs de seuil qui dépendent des niveaux est une solution pour éliminer un bruit coloré tout en étant plus adapté aux caractéristiques des signaux.

4.1 Seuillage global d'un bruit coloré :

Une représentation graphique de la GCV et de EQM (Fig 4.6) indique que la GCV ne permet plus de fournir une bonne approximation du seuil à minimum de risque lorsque le bruit est coloré. La raison est due à la non stationnarité du bruit dans le domaine des ondelettes. Nous avons un niveau de bruit σ qui dépend des coefficients ; en conséquence, la formule du SURE n'est plus valide. Et puisque les propriétés asymptotiques de la GCV sont basées sur cette estimation non-biaisée (SURE), le succès de la GCV n'est plus assuré.

A partir de l'étude de la matrice de corrélation des coefficients en ondelettes bruités on constate que si un bruit est non corrélé et stationnaire il reste non corrélé et stationnaire après une transformation en ondelettes orthogonale et si il est ni stationnaire ni blanc, sa transformée en ondelette va être ni blanche ni stationnaire.

La GCV atteint un seuil optimal si le nombre de coefficients tend vers l'infini sans nécessiter que les coefficients en ondelettes soient non corrélés. Par contre pour le seuil SURE, on a besoin que le bruit soit stationnaire dans le domaine des ondelettes. Mais dans ce cas trouver le seuil minimum EQM ne sera pas aussi utile car par intuition, on sait que plus les coefficients sont bruités plus le seuil doit être élevé. Si la quantité du bruit dépend des coefficients, il est difficile de l'éliminer par un seuil unique. Nous présentons dans la figure

(Fig 4.5) un exemple de reconstruction de coefficients après seuillage avec un seuil minimum EQM. Ce seuil paraît trop petit et ne peut enlever le bruit aux fines échelles. La plupart du bruit considéré est concentré dans cette région du fait qu'il est généré par un filtre passe haut correspondant aux fines échelles.

Si on arrivait à trouver la matrice de covariance de bruit Q alors il est possible de calculer la matrice de covariance S dans le domaine des ondelettes en utilisant (2.1). On pourrait ainsi normaliser les coefficients d'ondelette par :

$$w_i^{\text{norm}} = w_i / \sqrt{S_{ii}}$$

La procédure du seuillage par un seuil globale minimum MSE serait plus efficace dans l'élimination des coefficients bruités .

Malheureusement, en pratique on ne connaît pas la structure de corrélation de l'entrée et on compte sur le caractère multirésolution de la transformation en ondelettes. Celle-ci permet de prouver, en supposant que le bruit est stationnaire, que la variance des coefficients d'ondelettes de ce bruit à la position k et à la résolution i , $E(w_{i,k})^2$, ne dépend que de cet résolution i .

Dans le cas bidimensionnel, dans le plan des ondelettes, le niveau de bruit dépend aussi de l'orientation (horizontale, verticale, diagonale).

On montre qu'un bruit corrélé stationnaire est stationnaire pour chaque niveau, et comme la stationnarité est une condition nécessaire pour le succès de la GCV pour l'estimation d'un seuil optimal, le choix d'un seuil différent pour chaque niveau apparaît comme une solution.

Preuve :

Soit la corrélation du bruit au 1^{er} niveau de la transformation en ondelette comme suit :

$$E\omega_{1,k}\omega_{1,l} = \sum_u \sum_v \tilde{g}_{u-2k} \tilde{g}_{v-2l} E\eta_u \eta_v$$

On a : $E\eta_u \eta_v = Q_{u,v} = q_{|u-v|}$.

$$E\omega_{1,k}\omega_{1,l} = \sum_u \sum_v \tilde{g}_{u-2k} \tilde{g}_{v-2l} q_{|u-v|}.$$

en remplaçant $m=u-2k$ et $n=v-2l$ alors on obtient :

$$E\omega_{1,k}\omega_{1,l} = \sum_m \sum_n \tilde{g}_m \tilde{g}_n q_{|2(k-l)+m-n|}$$

De la dernière équation , on s'aperçoit que pour tout entier r , on a :

$$E\omega_{1,k+r}\omega_{1,l+r} = E\omega_{1,k}\omega_{1,l}$$

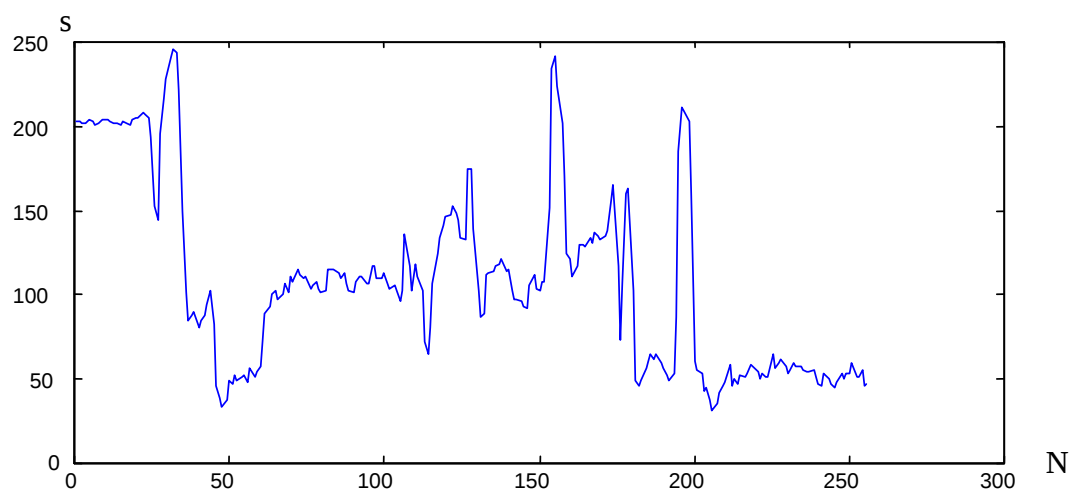


Fig 4.1 Le signal original

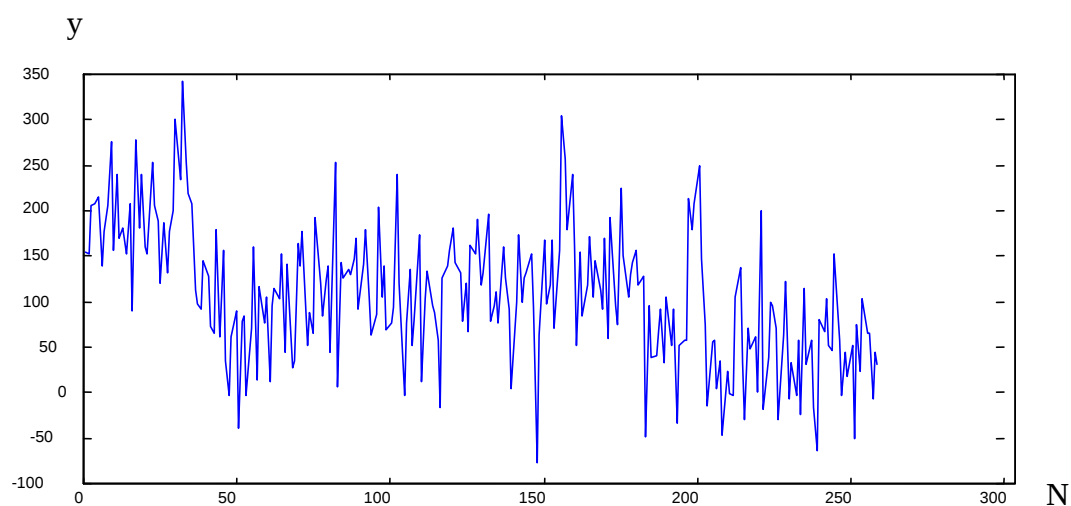


Fig 4.2 Le signal bruité avec un bruit coloré

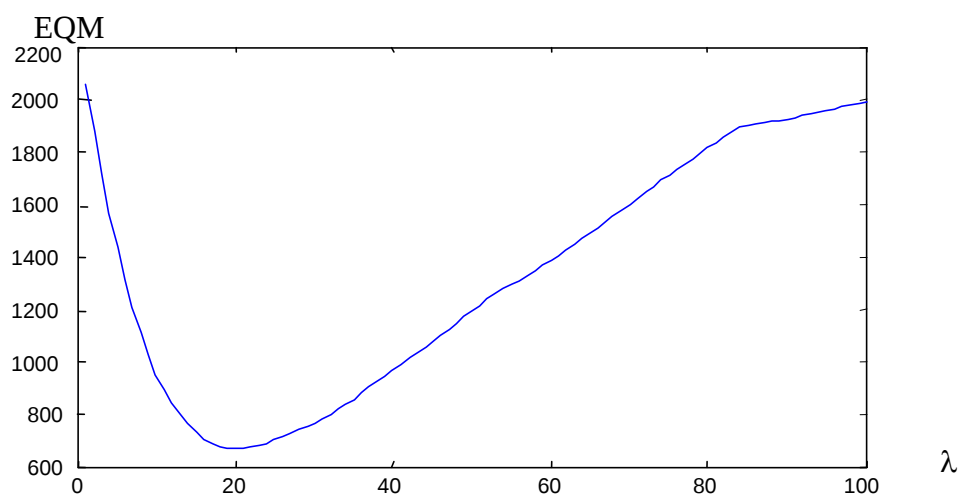


Fig 4.3 L'EQM de la fonction résultante en fonction du seuil

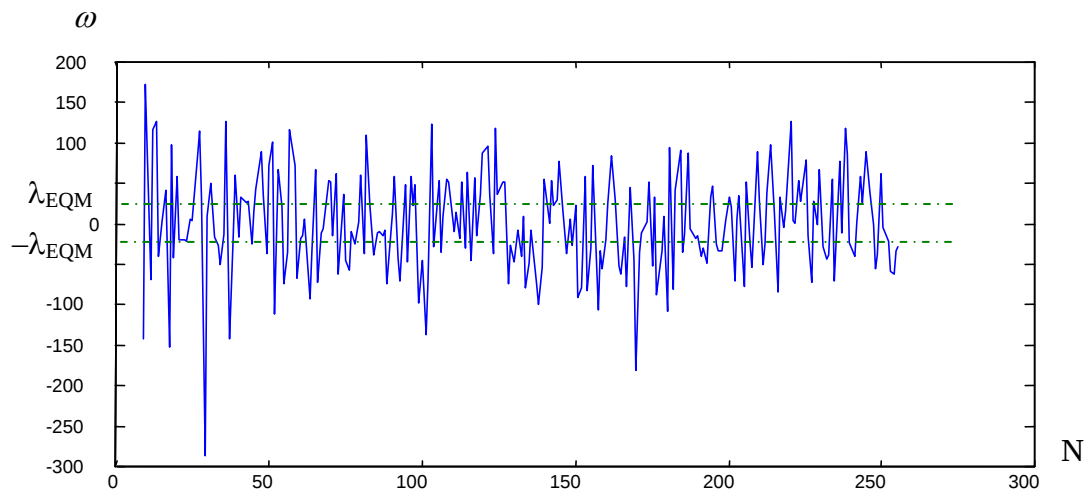


Fig 4.4 TO du signal bruité avec un bruit coloré

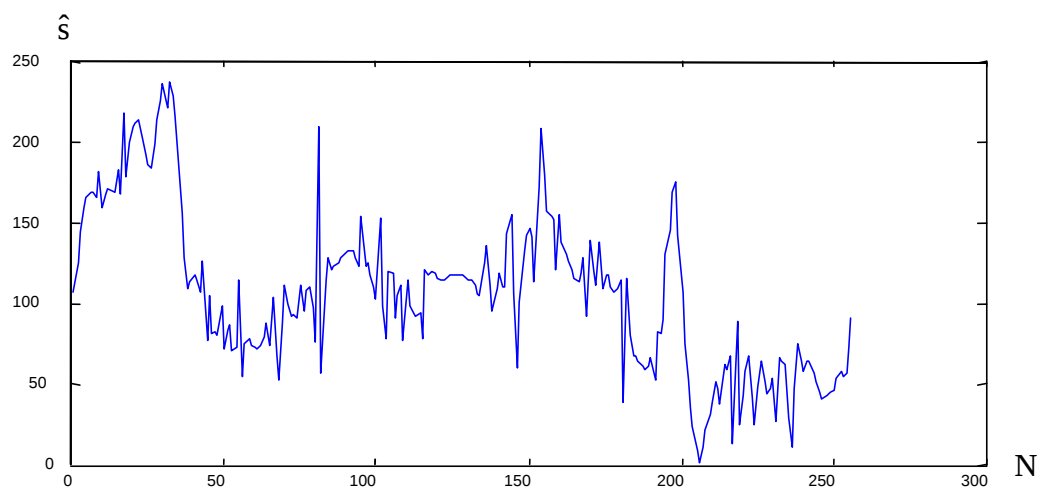
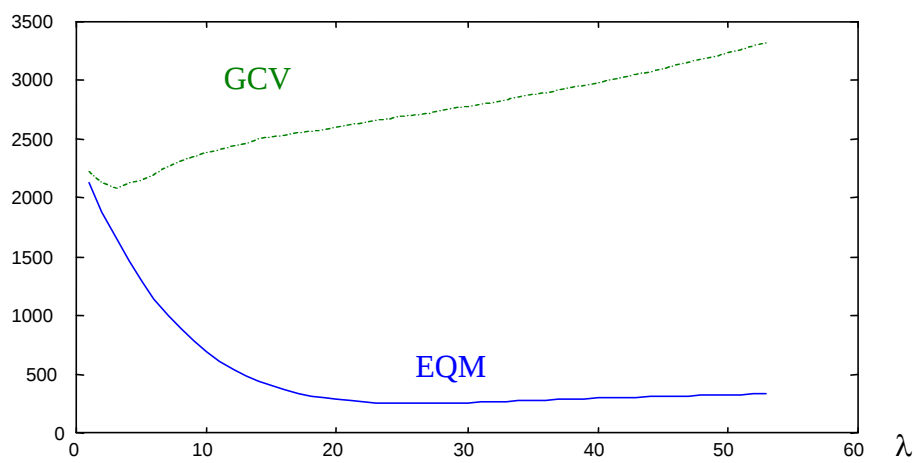


Fig 4.5 Le signal débruité avec un seuil qui minimise l'EQM

Fig 4.6 GCV (trait discontinu) et EQM (trait continu)
d'un signal bruité avec un bruit coloré

En particulier, on a :

$$E\omega_{l,k+r}^2 = E\omega_{l,k}^2 = \sigma_1^2.$$

On peut répéter la même procédure pour les niveaux suivant.

On peut en conclure que ω est stationnaire dans un même niveau.

4.2 Seuillage Universel dépendant des niveaux

Johnstone [Jon94] a proposé, dans le cas de signaux bruités par un bruit corrélé, un seuillage dépendant des niveaux, avec un seuil $T_i = \sigma_i \sqrt{2\log(N)}$ où σ_i est l'écart type du bruit au niveau i et N le nombre d'échantillons du signal de départ. Il a été démontré que le risque de ce seuillage est à un facteur $(2\log(N)+1)$ du risque atteint par un seuillage optimal :

$$R(\hat{s}, s) \leq (2\log(N) + 1)(\bar{\sigma}^2 + R(\text{optimal}));$$

$$\text{où } \bar{\sigma} = \frac{1}{N} \sum_i \sigma_i.$$

4.3 Extention au cas SURE et GCV

L'erreur quadratique moyenne est :

$$R(\lambda) = \sum_{i,c} \frac{N_i^c}{N} R_i^c(\lambda)$$

$$\text{Avec : } R_i^c(\lambda) = \frac{1}{N_i^c} \cdot \left\| w_{i\lambda}^c - v_i^c \right\|^2.$$

Où N_i^c est le nombre de coefficients dans le niveau i et orientation c et $R_i^c(\lambda)$ est le risque associé au niveau j et l'orientation c .

Puisque dans l'expression (5.1) tous les termes sont positifs, la minimisation de $R(\lambda)$ est équivalente à la minimisation des $R_i^c(\lambda)$ pour tout c et i .

De la même manière on a :

$$\text{SURE}_i^c = F_i^c(\lambda) + \sigma_i^2 \cdot \frac{N_i^c - 2N_{i0}^c}{N_i^c}.$$

$$\text{Où : } F_j^c(\lambda) = \frac{1}{N_i^c} \left\| w_{i\lambda}^c - w_i^c \right\|^2.$$

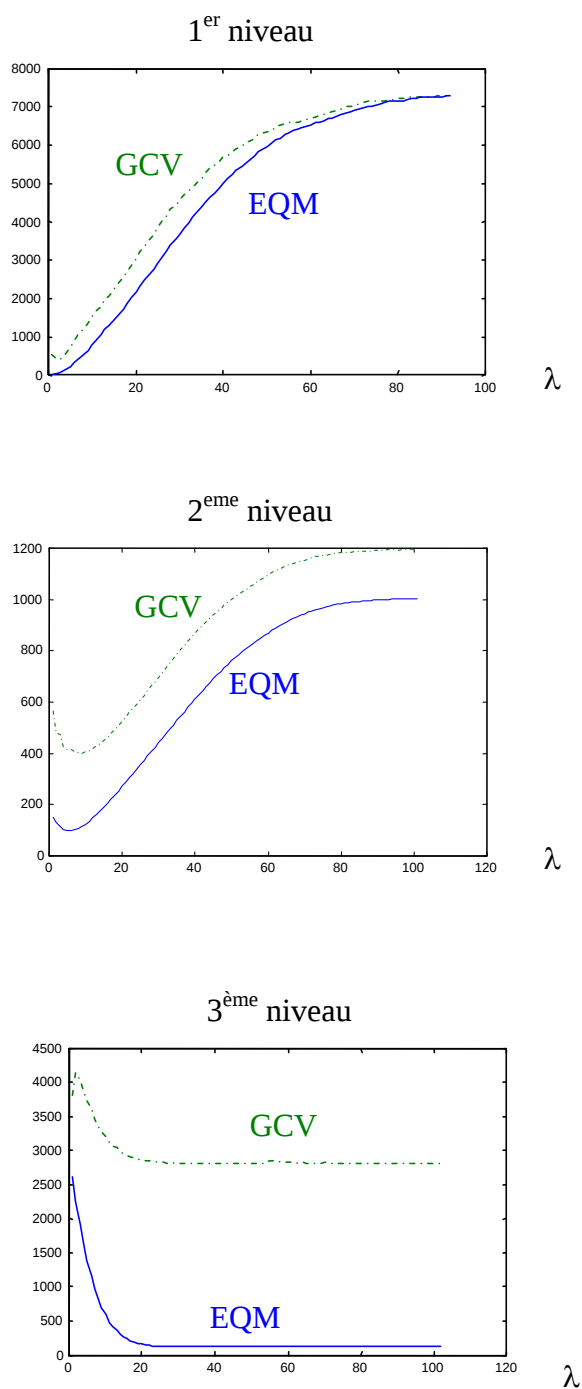


Fig 4.7 GCV et EQM en fonction du seuil pour les niveaux 1, 2 et 3 (le seuil est différent d'un niveau à un autre), la transformation en ondelette est avec décimation

Cet algorithme ainsi que le seuillage universel dépendant des niveaux exigent une estimation explicite du niveau de bruit à chaque échelle, nous utilisons l'estimation suivante [Joh94] :

$$\sigma_i = \text{MAD}(w_{i,k}) / 0.6745.$$

où MAD est l'écart absolu médian.

Ceci nous mène à une estimation adaptative telle que :

$$\text{GCV}_i^c(\lambda) = \frac{\frac{1}{N_i^c} \left\| \mathbf{w}_i^c - \mathbf{w}_{i\lambda}^c \right\|^2}{\left(\frac{N_{i0}^c}{N_i^c} \right)^2}$$

Le minimum de cette fonction est asymptotiquement optimal pour l'estimation à minimum de risque pour le niveau i et l'orientation c .

L'utilisation de la méthode GCV pour des signaux affectés par un bruit corrélé est motivée par le fait que les propriétés de la GCV ne sont pas affectées par le bruit corrélé directement.

Le bruit corrélé affecte seulement l'algorithme à travers la distribution des coefficients en ondelettes qui devient différente. Grâce à la structure multirésolution d'une transformation en ondelette, on est sûre que les coefficients ont une distribution identique dans un même niveau.

Le fait que la GCV ne nécessite pas une estimation du niveau du bruit à l'échelle i et l'orientation c est particulièrement intéressant car, dans le cas du bruit blanc, le niveau de bruit est le même à toutes les échelles. Il peut donc être estimé à partir des coefficients des fines échelles (cette résolution contient généralement peu de coefficients importants). Par contre, si le bruit est coloré ou si la transformation n'est pas orthogonale, on aura besoin d'une estimation à chaque niveau et chaque direction pour les larges échelles qui contiennent plusieurs coefficients de grande valeur.

4.4 Transformation non orthogonale

Si le bruit est non corrélé et stationnaire mais la transformation est biorthogonale, l'équation (2.1) nous apprend que les coefficients en ondelette sont corrélés et non stationnaires. On peut dans ce cas soit normaliser les coefficients soit appliquer un seuillage dépendant de l'échelle. Cette dernière solution a l'avantage d'adaptivité par rapport au signal et reste valable même si le bruit est coloré.

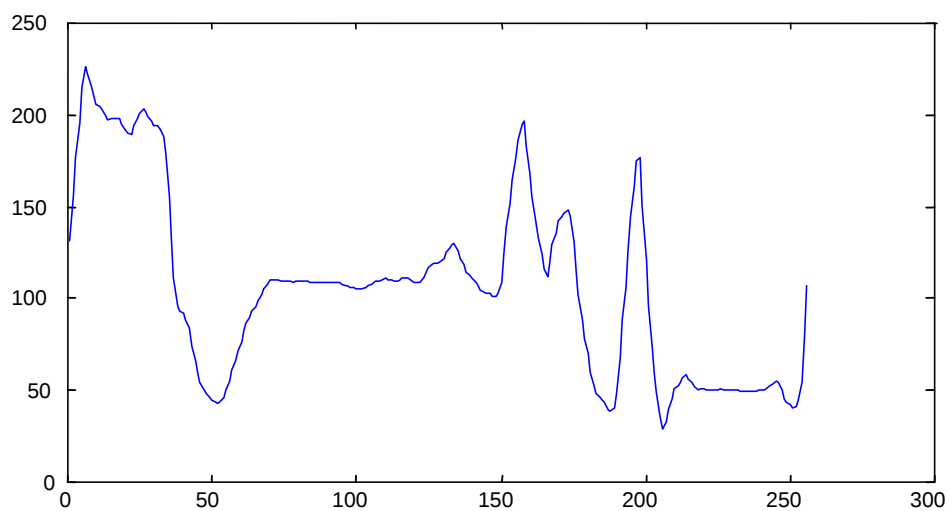


Fig 4.8 Résultat du seuillage universel dépendant des niveaux.

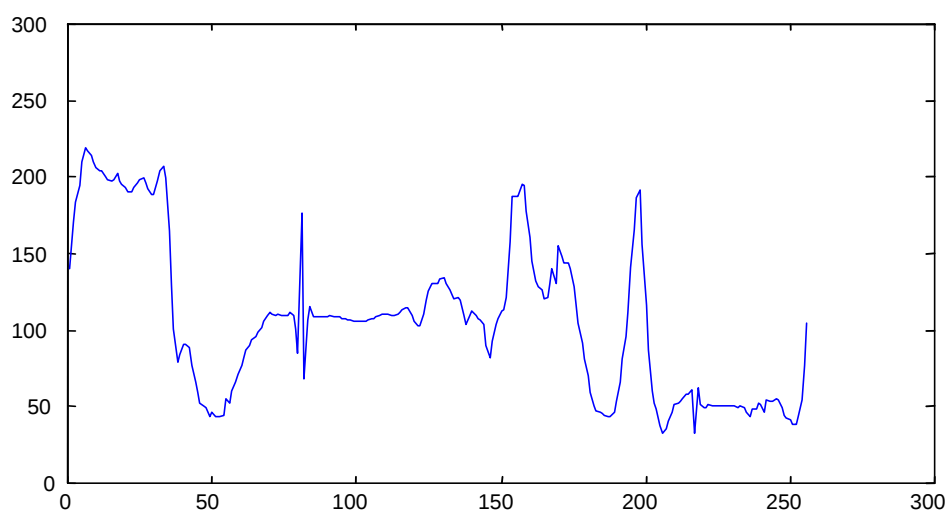


Fig 4.9 Résultat du seuillage SURE dépendant des niveaux.

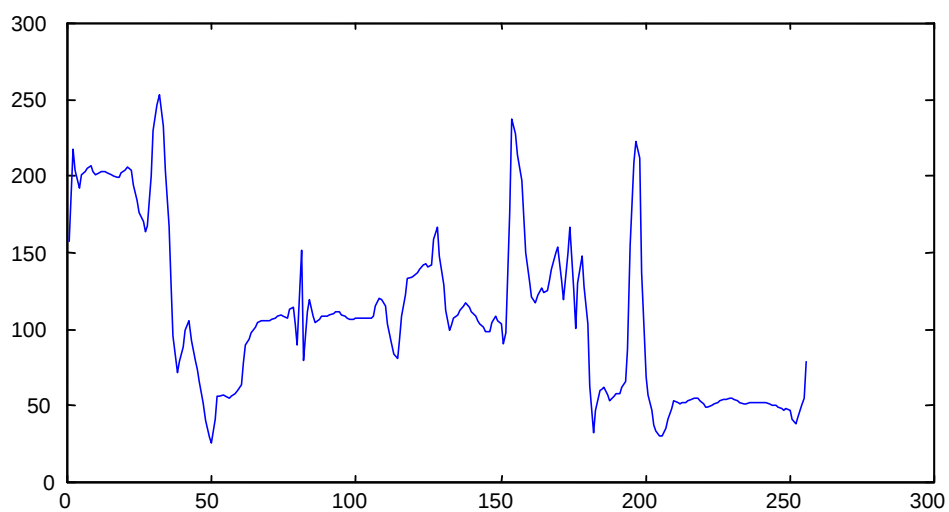


Fig 4.10 Résultat du seuillage GCV dépendant des niveaux.

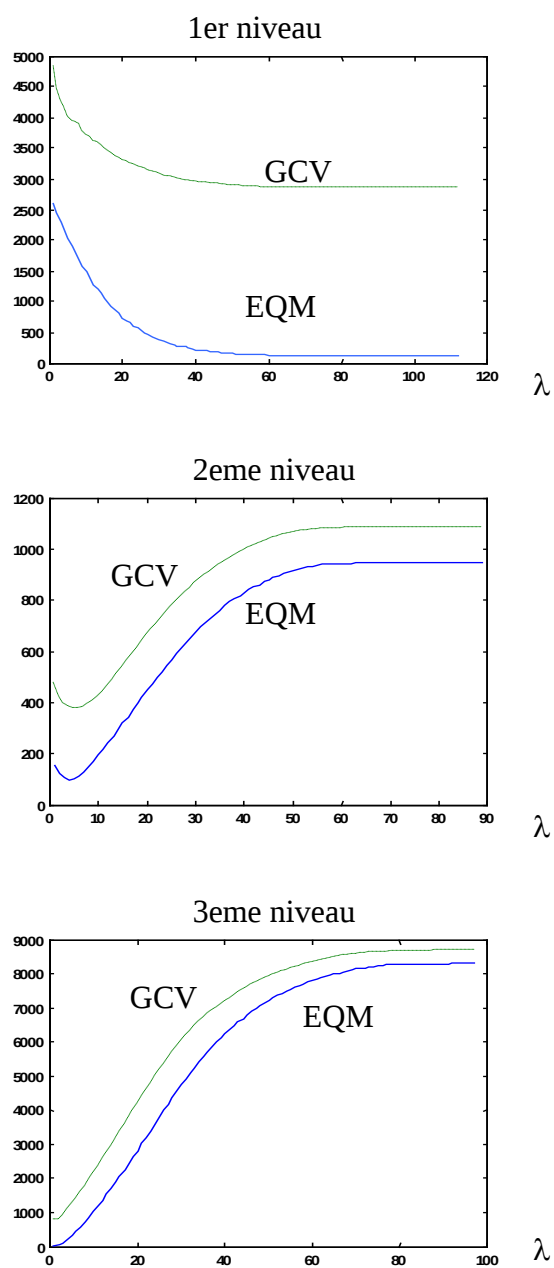


Fig 4.11 GCV et EQM en fonction du seuil pour les niveaux 1, 2 et 3 (le seuil est différent d'un niveau à un autre) dans le cas d'une transformation en ondelette sans décimation

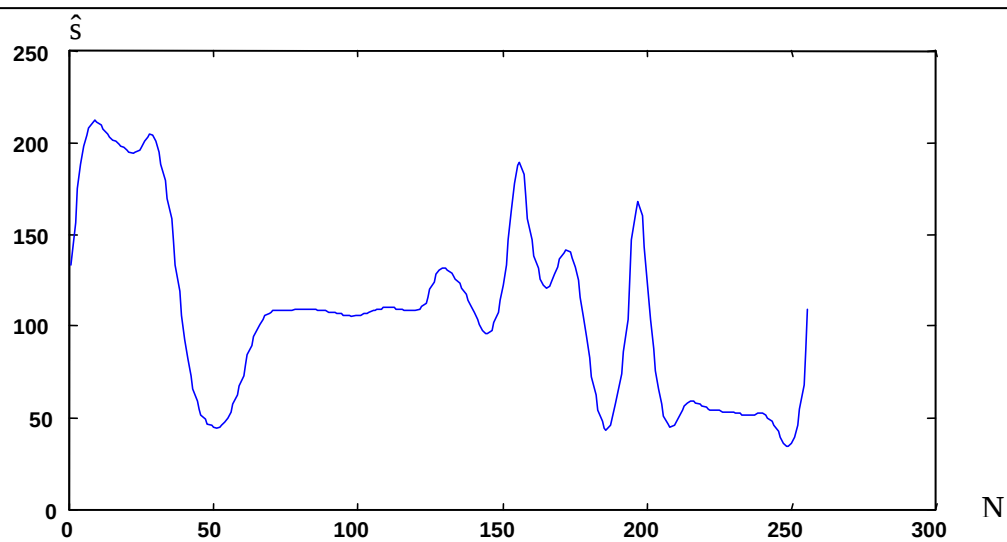


Fig 4.12 Résultat du seuillage universel dépendant des niveaux avec une transformation sans décimation

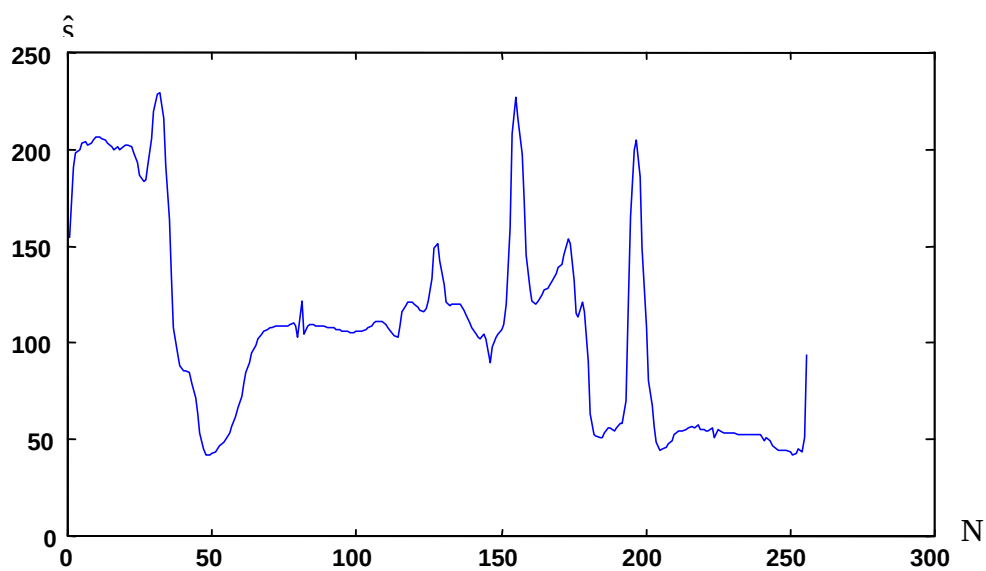


Fig 4.13 Résultat du seuillage SURE dépendant des niveaux avec une transformation sans décimation

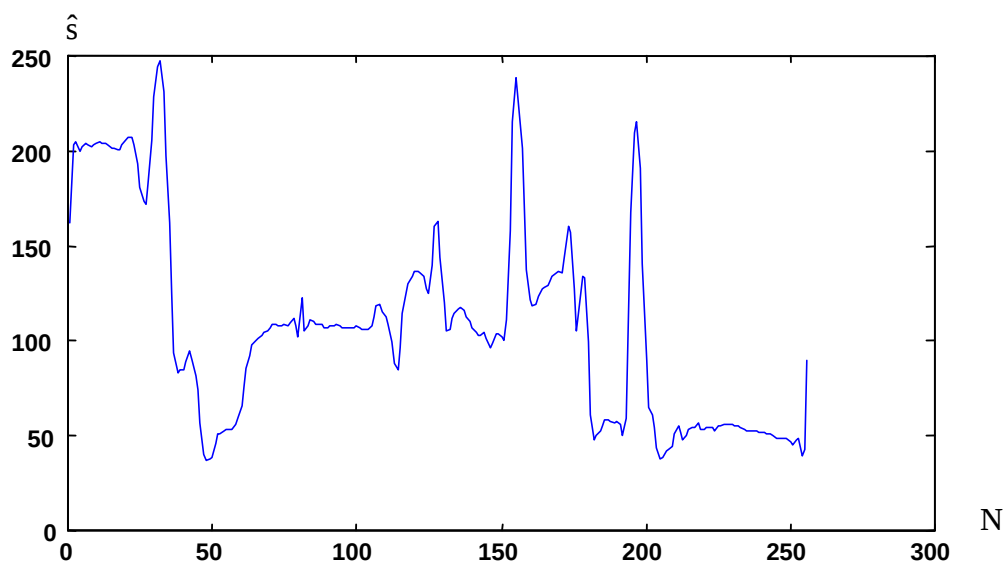


Fig 4.14 Résultat du seuillage GCV dépendant des niveaux avec une transformation sans décimation

4.5 Adaptivité pour chaque échelle

Même pour un bruit blanc et une transformation orthogonale, un seuillage dépendant du niveau peut être intéressant, car le seuil optimal ne dépend pas seulement du niveau de bruit, mais aussi des caractéristiques du signal. Ces caractéristiques peuvent être différentes à des niveaux différents. Les larges échelles exhibent une importante proportion de caractéristiques de signaux. La présence de larges coefficients force à ce que le seuil soit de faible valeur, même si le niveau du bruit est le même dans tous les niveaux le seuil à minimum de risque aux larges échelles tend à être petit. Si l'algorithme cherche un seuil global, il doit y avoir un compromis : les fines échelles vont sûrement trouver que le seuil est trop petit se qui se répercute en sortie par des structures bruitées « blips ».

4.6 Transformation en ondelette non décimée :

La TO non décimée a une complexité de calcul de $O(N \log(N))$ celle-ci est plus coûteuse que la TO avec décimation d'un facteur $\log(N)$. D'un autre côté, la TO sans décimation a quelques avantages.

- 1- Cette TO redondante est extensible au cas où le nombre de données n'est pas une puissance de deux.
- 2- La TO non décimée est redondante et génère un nombre égal de coefficients en ondelette à toutes les échelles. En principe, ceci facilite l'utilisation de la GCV dans les larges échelles, seulement la grande corrélation entre les coefficients réduit partiellement le bénéfice de cet avantage.
- 3- Il est facile de prouver que la TO redondante d'un bruit stationnaire est stationnaire à chaque échelle.
- 4- Contrairement à la TO décimée, la TO redondante est invariante par translation. Elle contient les transformations en ondelette de toutes les translations du signal d'entrée.

Si on définit l'opérateur de translation par S , tel que pour un signal y de taille N on a :

$$z = Sy \Leftrightarrow \begin{cases} z_k = y_{k-1} & \text{si } k = 2, \dots, N \\ z_1 = y_N \end{cases}$$

d'où la transformation en ondelette redondante contient tous les coefficients tel que :

$$w_n = \tilde{W} S^n y \quad n=1, \dots, N.$$

où $\tilde{W}$ est la transformation en ondelette avec décimation.

- 5- A chaque fois qu'on effectue une étape de transformation en ondelette inverse, on prend la moyenne des reconstructions des deux schémas de décomposition de l'échelle précédente. Ceci veut dire que pour réaliser la reconstruction globale, on calcule la moyenne des reconstructions des différentes transformations en ondelette des translations du signal d'entrée.

$$y_{\text{recon}} = \frac{1}{N} \sum_{n=1}^N S^{-n} W w_n.$$

$W = \tilde{W}^{-1}$ est la transformation en ondelette inverse, S^{-n} est l'opérateur de translation inverse

Ce qui cause évidemment un lissage de niveau du signal reconstruit.

Nous avons montré dans ce chapitre que la méthode GCV, pour l'estimation de l'EQM, peut être étendue aux signaux affectés par un bruit coloré et aux transformations non-orthogonales, en utilisant un seuillage dépendant des échelles dans le domaine des ondelettes. Car les conditions nécessaires pour la validité de la méthode GCV (stationnarité du bruit) sont assurées à chaque échelle. Cette solution a l'avantage supplémentaire de faire aboutir, dans tous les cas, à un résultat plus adaptatif aux caractéristiques du bruit et du signal.



Fig 5.1 L'image originale



Fig 5.2 Image bruitée (bruit blanc additif) à 10db

5.1. Evaluation qualitative

Nous proposons de tester l'approche d'élimination de bruit par décomposition en ondelette sur une image bruitée artificiellement avec un bruit blanc gaussien additif, le rapport signal sur bruit de l'image altérée est de 10 db. La décomposition a été réalisée jusqu'au 4^{ème} niveau. Pour avoir une idée sur les possibilités des approches classiques, nous commençons par présenter les résultats après application du filtre médian, et du filtre de Wiener, (Fig 5.3 et Fig 5.4) tous deux ont une fenêtre d'analyse de 3×3 . Les résultats sont globalement très proches. Une grande quantité de bruit persiste encore.

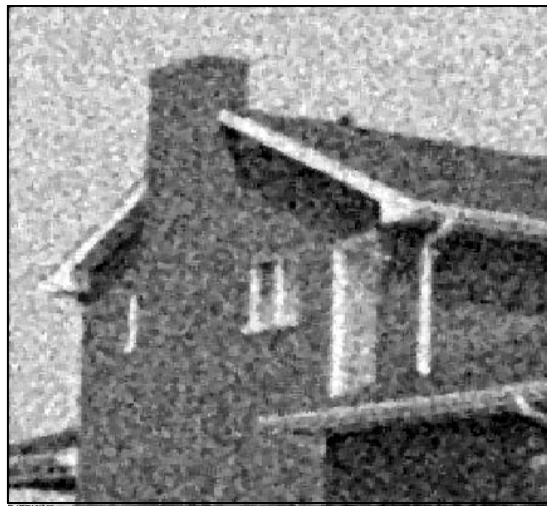


Fig 5.3. Résultat du filtre médian appliqué à l'image de la Fig 5.2



Fig 5.4. Résultat du filtre de Wiener appliqué à l'image de la Fig 5.2



Fig 5.5 Résultat du seuillage de Fig 5.2 avec un seuil universel global (SNR=19.84)



Fig 5.6 Résultat du seuillage de Fig 5.2 avec un seuil SURE global (SNR=18.95)



Fig 5.7 Résultat du seuillage de Fig 5.2 avec un seuil GCV global (SNR=19.10)



Fig 5.8 Résultat du seuillage de Fig 5.2 avec un seuil universel dépendant des niveaux (SNR=17.64)



Fig 5.9 Résultat du seuillage de Fig 5.2 avec un seuil SURE dépendant des niveaux (SNR=20.12)



Fig 5.10 Résultat du seuillage de Fig 5.2 avec un seuil GCV dépendant des niveaux (SNR=19.92)



Fig 5.11 Résultat du seuillage de Fig 5.2 avec un seuil universel dépendant des niveaux dans le cas où la transformation en ondelettes est sans décimation (SNR=16.14)



Fig 5.12 Résultat du seuillage de Fig 5.2 avec un seuil SURE dépendant des niveaux dans le cas où la transformation en ondelettes est sans décimation (SNR=21.39)



Fig 5.13 Résultat du seuillage de Fig 5.2 avec un seuil GCV dépendant des niveaux dans le cas où la transformation en ondelettes est sans décimation (SNR=21.18)

Dans le cas d'une image affectée par un bruit blanc, les résultats du seuillage global, malgré une élimination d'une grande partie du bruit, sont trop lissés, les contours sont mal préservés, surtout pour le seuil universel (Fig 5.5) où on arrive à peine à reconnaître certaines structures. Pour un seuillage avec le seuil SURE et GCV (Fig 5.6 et Fig 5.7), le lissage est moins important mais il reste inacceptable. En plus, on constate l'apparition de petites structures gênantes et qui pourraient même induire à une mauvaise interprétation des images.

L'application d'un seuillage dépendant des niveaux a des conséquences différentes selon le seuil choisi. Dans le cas du seuil universel, dépendant des niveaux (Fig 5.8), le lissage est encore fort malgré une légère amélioration dans la préservation des contours. Dans le cas du seuil SURE et GCV (Fig 5.9 et Fig 5.10), les contours apparaissent d'une manière plus nette, l'inconvénient est que les structures gênantes sont présentes avec une plus grande dimension. L'apport d'une transformation sans décimation, n'est pas très ressenti dans le cas du seuil universel (Fig 5.11). Mais, dans le cas des seuils se basant sur la minimisation de l'erreur quadratique moyenne (Fig 5.12 et Fig 5.13), cette transformation invariante par translation permet de se débarrasser des fausses structures gênantes sans réduire la qualité des contours déjà atteinte.



Fig 5.14 Image bruitée avec un bruit coloré additif (SNR = 10db)

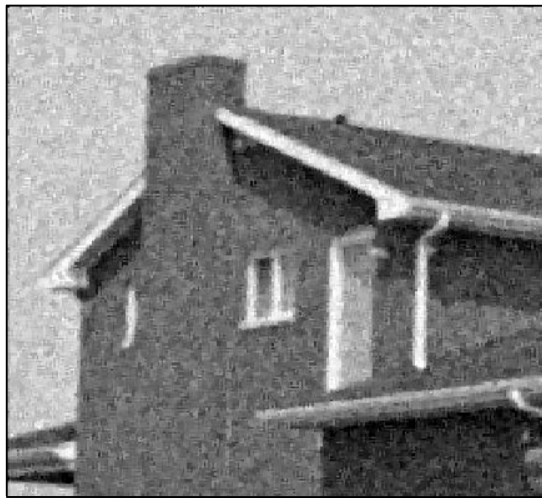


Fig 5.15 Résultat du filtre médian appliqué à Fig 5.14
(bruit coloré)



Fig 5.16 Résultat du filtre de Wiener appliqué à Fig 5.14
(bruit coloré)



Fig 5.17 Résultat du seuillage de Fig 5.14 avec un seuil universel global (SNR=17.86)



Fig 5.18 Résultat du seuillage de Fig 5.14 avec un seuil SURE global (SNR=19.18)

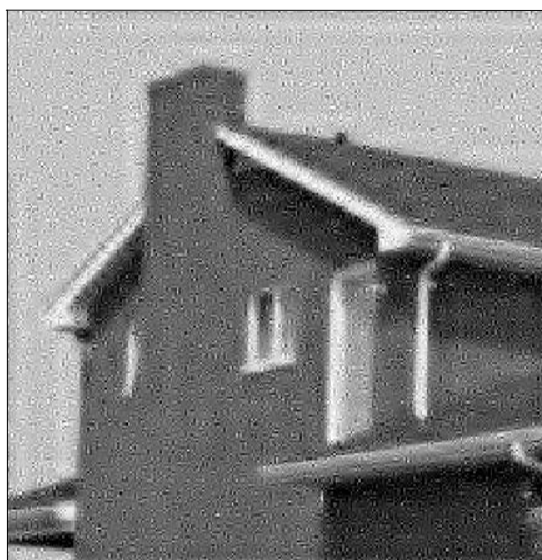


Fig 5.19 Résultat du seuillage de Fig 5.14 avec un seuil GCV global (SNR=15.48)



Fig 5.20 Résultat du seuillage de Fig 5.14 avec un seuil universel dépendant des niveaux (SNR=18.20)



Fig 5.21 Résultat du seuillage de Fig 5.14 avec un seuil SURE dépendant des niveaux (SNR=22)



Fig 5.22 Résultat du seuillage de Fig 5.14 avec un seuil GCV dépendant des niveaux (SNR=22.97)



Fig 5.23 Résultat du seuillage de Fig 5.14 avec un seuil universel dépendant des niveaux dans le cas où la transformation et sans décimation (SNR=18.42)



Fig 5.24 Résultat du seuillage de Fig 5.14 avec un seuil SURE dépendant des niveaux dans le cas où la transformation et sans décimation (SNR=23)



Fig 5.25 Résultat du seuillage de Fig 5.14 avec un seuil universel dépendant des niveaux dans le cas où la transformation et sans décimation (SNR=26)

Nous suivrons le même cheminement pour l'évaluation des résultats de notre approche sur la même image bruitée cette fois-ci avec un bruit coloré (Fig 5.14). Le bruit coloré a été obtenu en filtrant le bruit blanc gaussien avec un filtre passe haut. La décomposition se fait jusqu'au 3^{ème} niveau. Celle-ci est suffisante du fait que le bruit est concentré dans les échelles les plus fines.

Comme précédemment, nous commençons par présenter les résultats des méthodes classiques, filtre médian et filtre de Wiener . Cette fois-ci les résultats sont différents : Le filtre médian n'arrive pas à éliminer le bruit correctement (Fig 5.15) mais le résultat du filtre de Wiener (Fig 5.16) est nettement meilleur bien que le bruit n'a pas été totalement supprimé. Les méthodes se basant sur le seuillage des coefficients en ondelettes avec un seuil global universel et SURE (Fig 5.17 et Fig 5.18) donnent des résultats lissés avec des structures mal représentées ; on voit qu'une grande partie de l'information a été éliminée avec le bruit. Quant au résultat d'un seuillage avec un seuil GCV global (Fig 5.19), il reste trop bruité. Ceci vient du fait que la méthode GCV échoue dans l'estimation de l'erreur quadratique moyenne

L'apport d'un seuillage dépendant des niveaux ne se fait ressentir que pour les méthodes où le seuil minimise l'EQM, car pour le cas du seuil universel l'amélioration n'est pas perceptible (Fig 5.20). Pour le seuil SURE une amélioration considérable est constatée dans le dessin des contours sauf que certaines informations ne sont pas représentées et un effet d'image double apparaît (Fig 5.21). L'image débruitée avec le seuil GCV dépendant des niveaux (Fig 5.22) est nettement meilleur d'une part par rapport au cas du seuillage GCV global et d'autre part par rapport au seuil SURE dépendant des niveaux, une élimination quasi-totale du bruit est obtenue avec une bonne préservation des contours.

L'application d'une décomposition invariante par translation améliore les résultats en supprimant l'effet d'escalier des contours, sans détériorer les résultats atteints dans le cas sans décimation, il est clair que les seuils se basant sur la minimisation de l'erreur quadratique moyenne sont ceux qui atteignent les meilleurs résultats (Fig 5.24 et Fig 5.25), avec un net avantage pour le seuil GCV.

5.2. Evaluation quantitative

L'évaluation quantitative de la réduction du bruit se fait par le biais de l'estimation du rapport signal sur bruit (SNR) dans l'image résultante. On définit ce rapport de la manière suivante :

$$SNR = 10 \log_{10} \left(\frac{\text{variance}(\text{image})}{\text{variance}(\text{bruit restant})} \right);$$

Le SNR traduit le rapport entre l'énergie signal sur l'énergie du bruit, une augmentation de cette mesure est synonyme d'une amélioration du résultat en terme d'élimination du bruit.

Le SNR des résultats des seuillages appliqués à une image bruitée avec un bruit blanc gaussien à un taux de 10db sont résumés dans le tableau suivant :

	Seuil Universel	Seuil SURE	Seuil GCV
Un seuil global appliqué à tous les coefficients	16.8469	18.9549	19.1028
Seuillage dépendant des niveaux avec une transformation avec décimation	17.6487	20.1233	19.9250
Seuillage dépendant des niveaux avec une transformation sans décimation	16.1460	21.3956	21.1825

Tab 1 : Rapport signal sur bruit
(Cas d'un seuillage global et d'un bruit blanc)

Le SNR des résultats des seuillages appliqués à une image bruitée par un bruit coloré à un taux de 10db sont les suivants :

	Seuil universel	Seuil SURE	Seuil GCV
Un seuil unique pour tous les niveaux	17.86	19.18	15.48
Seuillage dépendant des niveaux avec une transformation avec décimation	18.20	22	23.13
Seuillage dépendant des niveaux avec une transformation sans décimation	18.42	22.97	25.92

Tab 2 : Rapport signal sur bruit
(Cas d'un seuillage global et d'un bruit coloré)

On procède aussi à l'évaluation de l'erreur normalisée entre l'image sans bruit et l'image débruitée par seuillage avec un seuil λ . Celle-ci exprimée en % est définie comme suit :

$$\text{Erreur}_{\lambda} = \frac{\sum_i \sum_j (\hat{s}_{i,j,\lambda} - s_{i,j})^2}{\sum_i \sum_j s_{i,j}^2}.$$

Les erreurs normalisées des résultats des seuillages appliqués à une image bruitée avec un bruit blanc gaussien sont présentées dans le tableau 3. L'erreur de l'image avant seuillage est de 1.002 %. Il est clair qu'une diminution de cette erreur reflète un rapprochement entre l'image débruitée et l'image sans bruit.

	Seuil Universel	Seuil SURE	Seuil GCV
Un seuil unique pour tous les niveaux	2.055	1.265	1.223
Seuillage dépendant des niveaux avec une transformation avec décimation	1.709	0.967	1.012
Seuillage dépendant des niveaux avec une transformation sans décimation	2.135	0.755	0.779

tab 3 : Erreurs normalisées des images résultantes
(Cas d'un seuillage global et d'un bruit blanc)

	Seuil Universel	Seuil SURE	Seuil GCV
Un seuil unique pour tous les niveaux	1.624	1.200	2.788
Seuillage dépendant des niveaux avec une transformation avec décimation	1.503	0.627	0.331
Seuillage dépendant des niveaux avec une transformation sans décimation	1.429	0.500	0.253

tab 4 : Erreurs normalisées des images résultantes
(Cas d'un seuillage global et d'un bruit coloré)

En comparant entre les méthodes, on peut constater une concordance totale entre le rapport signal sur bruit et l'erreur normalisée. Ce qui nous mène à l'analyse suivante :

- Dans le cas d'une image affectée par bruit blanc gaussien, on peut remarquer d'après les deux mesures que les performances des méthodes se basant sur la minimisation de l'EQM sont supérieures à celles obtenues avec les méthodes se basant sur le seuil universel. On note que les méthodes SURE et GCV donnent des résultats qui sont très proches en terme de SNR et Erreur_λ . Un seuillage dépendant des niveaux apporte une amélioration aux différentes méthodes. A l'exception du seuil universel, l'amélioration est encore plus importante si on applique une transformation sans décimation.
- Dans le cas d'une image corrompue par un bruit coloré, on peut remarquer que les méthodes se basant sur un seuillage universel ou sur un seuillage utilisant un seuil global donnent des mesures de SNR et de Erreur_λ qui sont inférieures à celles obtenues avec le filtre médian ou le filtre de Wiener tableaux 7 et 8, tout particulièrement la méthode GCV avec un seuil global où on enregistre les plus mauvaises performances. L'application d'un seuillage dépendant des niveaux contribue dans l'amélioration des résultats en particulier la méthode qui utilise un seuil GCV. L'application d'une transformation en ondelette sans décimation est bénéfique, ceci est prouvé par l'augmentation du rapport signal sur bruit (SNR) et la diminution de l'erreur normalisée (Erreur_λ). On remarque aussi un écart entre les performances des méthodes se basant sur la minimisation de L'EQM (le seuil GCV donne des résultats dont les performances sont supérieures à celles obtenues avec le seuil SURE).

Du fait que le rapport signal sur bruit (SNR) et l'erreur normalisée (Erreur_λ) sont calculés d'une manière globale par rapport à l'image, l'évolution de leurs valeurs ne peut être utilisée pour traduire une amélioration absolue de l'état de l'image. On doit utiliser d'autre mesure pour mieux évaluer l'effet du seuillage notamment mesurer la préservation des contours. Pour cela, nous complétons l'évaluation des performances par deux mesures :

La première est une mesure de corrélation définie comme suit :

$$\rho = \frac{\Gamma(s - \bar{s}, \hat{s} - \bar{\hat{s}})}{\sqrt{\Gamma(s - \bar{s}, s - \bar{s})\Gamma(\hat{s} - \bar{\hat{s}}, \hat{s} - \bar{\hat{s}})}};$$

où s : l'image sans bruit ;

$\hat{s}$: l'image résultat du seuillage ;

$\bar{s}, \bar{\hat{s}}$: respectivement moyenne de s et $\hat{s}$;

$$\text{et } \Gamma(s_1, s_2) = \sum_{i,j} s_1(i, j) s_2(i, j);$$

La seconde mesure est une mesure de préservation des contours :

$$\alpha = \frac{\Gamma(\nabla^2(s) - \overline{\nabla^2(s)}, \nabla^2(\hat{s}) - \overline{\nabla^2(\hat{s})})}{\sqrt{\Gamma(\nabla^2(s) - \overline{\nabla^2(s)}, \nabla^2(s) - \overline{\nabla^2(s)}) \Gamma(\nabla^2(\hat{s}) - \overline{\nabla^2(\hat{s})}, \nabla^2(\hat{s}) - \overline{\nabla^2(\hat{s})})}};$$

où $\nabla^2(s(i,j))$ est un résultat d'un filtrage passe haut de $s(i,j)$; obtenu par une approximation du Laplacien avec une fenêtre 3×3 [Kun 88]. Les mesures ρ et α seront proches de l'unité lorsque l'image filtrée est similaire à l'image référence

Ces mesures son résumées dans les tableaux 5 et 6 suivants :

Dans le cas d'un bruit blanc :

L'image est bruitée à 10db : ρ en % = 77.73 ; α en %=10.94 ;

		Seuillage par un seuil unique	Seuillage dépendant des niveaux	Seuillage dépendant des niveaux avec une T.O sans décimation
ρ en %	UNIV	93.15	94.35	91.93
	SURE	95.81	96.81	97.64
	GCV	95.97	96.67	97.56
α en %	UNIV	10.48	13	14.81
	SURE	11.48	18.42	32.26
	GCV	11.58	17.24	31.31

Tab 5 Les mesures de corrélation et de préservation de contours
(Cas d'un seuillage global et d'un bruit blanc)

Dans le cas d'un bruit coloré :

L'image est bruitée à 10db : ρ en % = 77.95 ; α en % = 9.37 ;

		Seuillage par un seuil unique	Seuillage dépendant des niveaux	Seuillage dépendant des niveaux avec une T.O sans décimation
ρ en %	UNIV	94.58	95	95.7
	SURE	96.03	97.95	98.4
	GCV	91.8	98.92	99.18
α en %	UNIV	7.71	9.86	12.64
	SURE	7.34	23.03	28.57
	GCV	9.17	29.63	43.35

Tab 6 Les mesures de corrélation et de préservation de contours
(Cas d'un seuillage global et d'un bruit coloré)

A titre comparatif on présente les résultats des mesures de performances effectuées sur les images filtrées par le filtre médian et filtre de Wiener dans le cas d'un bruit blanc tableau 7 et bruit coloré tableau 8.

	SNR	Erreur Normalisée Erreur_λ	Corrélation	Préservation des contours
Filtre médian	16.94	2.01	93.72	9.52
Filtre Wiener	17.02	1.97	93.90	14.53

Tab 7 Performances des filtres médian et de Wiener
(Cas d'un seuillage global et d'un bruit blanc)

	SNR	Erreur Normalisée Erreur_λ	Corrélation	Préservation des contours
Filtre médian	19.35	1.149	96.29	8.61
Filtre Wiener	20.03	0.985	96.83	12.62

Tab 8 Performances des filtres médian et de Wiener
(Cas d'un seuillage global et d'un bruit coloré)

D'après le tableau 5, on peut remarquer que le plus mauvais résultat, en terme de corrélation et de préservation de contours, a été obtenu en utilisant un seuil universel global. L'application d'un seuillage dépendant des niveaux améliore légèrement ces deux mesures dans le cas où la TO est avec décimation, cependant si celle-ci est sans décimation l'amélioration ne concerne que la mesure de préservation de contours α car la corrélation se détériore et atteint une valeur encore plus faible que celle obtenue avec un seuil global.

Les méthodes se basant sur la minimisation de EQM, SURE ou GCV, présentent des mesures de corrélation et de préservation des contours meilleurs en comparaison avec celles obtenues précédemment. Une amélioration de ces valeurs est enregistrée si on applique un seuillage dépendant des niveaux, spécialement la préservation des contours dans le cas des seuils SURE et GCV. Cette amélioration est encore plus sensible si la TO utilisée est sans décimation, on remarque aussi que les meilleurs résultats sont obtenus en appliquant un seuil SURE dépendant des niveaux avec une transformation en ondelette sans décimation.

Dans le cas d'un bruit coloré tableau 6, l'analyse des résultats présentés est plus simple. Le seuil universel présente globalement une mesure de corrélation ρ et de préservation de contours α inférieurs à celle atteinte par le seuil SURE ou le seuil GCV, ceci ne s'applique pas à la mesure de corrélation dans le cas du seuil GCV global. Les mesures de préservation de contours α des seuillages avec un seuil global sont inacceptables car elles sont inférieures à leurs valeurs avant filtrage. Pour cela, l'application d'un seuillage dépendant des niveaux ne peut qu'améliorer ces mesures, en effet on constate une nette augmentation des valeurs de ρ et de α spécialement dans les cas des seuils SURE et GCV. Là aussi une transformation sans décimation contribue dans l'évolution de ces mesures.

On propose de présenter l'effet du seuillage avec un seuil GCV sur des images bruitées avec un rapport signal sur bruit SNR qui augmente progressivement de -6db à 50db . L'image originale est celle présentée dans la figure Fig 5.1.

Cette procédure sera appliquée aux images affectées par un bruit blanc, en suite aux images corrompues par un bruit coloré.

Dans le cas où le bruit est blanc, le tableau 9 contient les mesures de rapport signal sur bruit SNR, de corrélation ρ , de préservation des contours α et d'erreur normalisée des images avant seuillage. On utilisera en premier un seuillage dépendant des niveaux avec une TO avec décimation, le tableau 10 renferme les résultats correspondants. Les évaluations sur les résultats d'un seuillage dépendant des niveaux avec une TO sans décimation sont présentées dans le tableau.11

SNR avant seuillage	ρ %	α %	Erreur %
50	99.99	99.61	$9.93 \cdot 10^{-4}$
40	99.96	96.23	$1.01 \cdot 10^{-2}$
30	99.67	74.72	$9.99 \cdot 10^{-2}$
20	96.91	33.41	$9.95 \cdot 10^{-1}$
12	84.22	13.83	$6.27 \cdot 10^{-1}$
9	74.15	9.77	1.25
6	61.56	6.86	2.49
3	48.32	4.79	4.98
0	36.30	3.32	9.93
-3	26.48	2.28	19.8
-6	18.97	1.54	39.5

Tab 9. Les performances avant seuillage (bruit blanc)

SNR avant seuillage	SNR après seuillage	ρ %	α %	Erreur %
50	36.20	99.92	97.78	$2.38 \cdot 10^{-2}$
40	35.18	99.9	94.72	$3.01 \cdot 10^{-2}$
30	30.87	99.73	81.53	$8.13 \cdot 10^{-2}$
20	25.47	99.08	51.66	$2.82 \cdot 10^{-1}$
12	21.08	97.45	22.46	$7.75 \cdot 10^{-1}$
9	19.56	96.36	15.24	1.10
6	17.96	94.72	6.53	1.58
3	16.58	92.78	3.92	2.18
0	14.79	89.39	2.24	3.30
-3	12.91	84.33	1.06	5.08
-6	10.63	76.47	0.67	8.59

Tab 10 Performances après seuillage dépendant des niveaux (TO avec décimation)

SNR avant seuillage	SNR après seuillage	ρ %	α %	Erreur %
50	36.68	99.93	98.36	$2.13 \cdot 10^{-2}$
40	35.95	99.92	96.35	$2.53 \cdot 10^{-2}$
30	31.93	99.79	86.62	$6.38 \cdot 10^{-2}$
20	26.5	99.28	64.47	$2.22 \cdot 10^{-1}$
12	22.16	98.03	37.57	$6.04 \cdot 10^{-1}$
9	20.56	97.12	28.51	$8.74 \cdot 10^{-1}$
6	19.03	95.87	18.63	1.24
3	17.51	94.10	12.07	1.17
0	15.87	91.39	4.95	2.57
-3	14.15	87.48	2.61	3.82
-6	12.09	81.22	1.66	6.15

Tab 11 Performance après seuillage dépendant des niveaux (TO sans décimation)

On constate que si l'image à un rapport signal sur bruit supérieur à 30 db l'application d'un seuil GCV dépendant des niveaux (avec ou sans décimation) conduit à un résultat ayant des performances inférieures de l'image avant seuillage. On remarque aussi que pratiquement dans tous les cas utilisation d'une TO sans décimation apporte une amélioration supplémentaire.

Dans le cas où le bruit est coloré, en comparant les différentes mesures avant seuillage (tableau 12) et après seuillage (tableau 13), pour le cas avec décimation, et (tableau 14) pour le cas sans décimation, on constate aussi que le seuillage réduit légèrement la qualité du résultat dès que le SNR de l'image à filtrer dépasse 30db. Le seuillage dépendant des niveaux avec un seuil GCV montre de bons résultats même si le signal est noyé dans le bruit (SNR négatif).

SNR avant seuillage	ρ %	α %	Erreur %
50	99.99	99.46	$9.93 \cdot 10^{-4}$
40	99.96	95.02	$9.93 \cdot 10^{-3}$
30	99.67	69.35	$9.93 \cdot 10^{-2}$
20	96.91	28.98	$9.93 \cdot 10^{-1}$
12	84.22	11.82	6.27
9	74.27	8.33	12.5
6	61.75	5.83	24.9
3	48.57	4.06	49.8
0	36.6	2.8	99.3
-3	26.82	1.91	198
-6	19.33	1.28	395

Tab 12 . Les performances avant seuillage (bruit coloré)

SNR avant seuillage	SNR après seuillage	ρ %	α %	Erreur %
50	36.33	99.92	97.77	$2.31 \cdot 10^{-2}$
40	35.37	99.9	94.45	$2.88 \cdot 10^{-2}$
30	31.51	99.77	80.53	$7 \cdot 10^{-2}$
20	27.43	99.41	52.51	$1.79 \cdot 10^{-1}$
12	25.08	98.99	31.66	$3.08 \cdot 10^{-1}$
9	24.67	98.89	29.16	$3.38 \cdot 10^{-1}$
6	24.02	98.71	25.57	$3.93 \cdot 10^{-1}$
3	23.05	98.39	21.71	$4.49 \cdot 10^{-1}$
0	22	97.94	14.25	$6.25 \cdot 10^{-1}$
-3	20.79	97.27	8.92	$8.26 \cdot 10^{-1}$
-6	19.88	96.62	5.65	$10.21 \cdot 10^{-1}$

Tab 13 Performances après seuillage (TO avec décimation)

SNR avant seuillage	SNR après seuillage	ρ %	α %	Erreur %
50	36.7	99.93	98.36	$2.12 \cdot 10^{-2}$
40	35.87	99.92	95.73	$2.57 \cdot 10^{-2}$
30	32.29	99.81	85.54	$5.85 \cdot 10^{-2}$
20	28.64	99.56	65.47	$1.35 \cdot 10^{-1}$
12	26.27	99.24	45.38	$2.34 \cdot 10^{-1}$
9	25.88	99.17	42.68	$2.5 \cdot 10^{-1}$
6	25.13	99.03	39.90	$3.01 \cdot 10^{-1}$
3	24.41	98.85	34.29	$3.5 \cdot 10^{-1}$
0	23.48	98.55	24.03	$4.45 \cdot 10^{-1}$
-3	22.03	98.09	14.77	$5.84 \cdot 10^{-1}$
-6	21.51	97.70	7.41	$7.01 \cdot 10^{-1}$

Tab 14 Performances après seuillage (TO sans décimation)

Conclusion

Il apparaît de l'évaluation visuel et de l'analyse des mesures effectuées sur les résultats des opérations de seuillage, dans le cas d'un bruit blanc ou d'un bruit coloré, que le seuil universel n'est pas adapté à l'élimination du bruit dans les images. Ceci est dû au lissage excessif des images provoquant une détérioration dans leurs qualités et une perte conséquente des informations qu'elles contenaient. L'application d'un seuillage dépendant des niveaux, en utilisant une TO avec ou sans décimation, bien qu'il améliore légèrement les résultats reste, dans ce cas, inefficace et insuffisant. Les seuils se basant sur la minimisation de l'erreur quadratique moyenne (EQM) ont permis d'obtenir de meilleurs résultats qui réunissent lissage du bruit et préservation des contours, toutefois une exception se présente et doit être signalée dans le cas du seuillage global d'une image bruitée avec un bruit coloré par un seuil GCV. L'apport d'un seuillage dépendant des niveaux est positif, car il permet d'améliorer les résultats, en particulier dans le cas d'un bruit coloré, en offrant plus d'adaptivité non seulement avec les caractéristiques du bruit mais aussi celles de l'image. L'utilisation d'une TO sans décimation s'est avérée bénéfique dans la plupart des cas car celle-ci participe dans la progression de la qualité des résultats en lissant les zones homogènes et en apportant une douceur aux transitions. Il faut souligner une différence fondamentale entre les seuils SURE et GCV qui consiste dans le fait que le seuil GCV ne nécessite aucune information sur le bruit contrairement au seuil SURE qui a besoin d'une estimation du niveau du bruit. Malgré ça le seuil GCV réussit à égaler le seuil SURE dans le cas du bruit blanc et même à le dépasser (si le seuillage étant dépendant des niveaux) dans le cas d'un bruit coloré. L'apport bénéfique d'une transformation en ondelette sans décimation pour le seuil GCV s'est avéré incontestable.

Conclusion générale

Bien que les méthodes classiques d'élimination de bruit se sont avérées satisfaisantes dans certaines applications, et ont l'avantage d'être simples dans leurs mises en œuvre, on constate certaines insuffisances dans des cas plus généraux où on doit concilier élimination de bruit et préservation de l'information contenue dans l'image. Ceci est lié au problème de choix de la taille de la fenêtre d'analyse. Pour cette raison, notre orientation vers des méthodes se basant sur la transformation en ondelettes semble naturelle vue que celle-ci offre une représentation temps-échelle de l'image à traiter permettant une action locale en espace et en fréquence.

On s'est axé sur le seuillage des coefficients en ondelettes pour ne retenir que ceux qui sont importants et permettent l'obtention par transformation inverse une image sans bruit. La question cruciale à résoudre était comment trouver le seuil qui sépare les coefficients significatifs à préserver des coefficients bruités à éliminer ?

Nous avons opté pour le seuil qui minimise l'erreur quadratique moyenne. Mais comme cette mesure est impossible à trouver car elle dépend de l'information sans bruit qui est inconnue, on doit passer par des estimateurs de l'erreur quadratique moyenne. Deux estimateurs ont été utilisés : l'estimateur de Stein non biaisé qui nécessite la valeur du niveau du bruit et la validation croisée généralisée qui ne nécessite aucune information sur le bruit.

Cette méthode a été appliquée au cas classique d'une image altérée par un bruit blanc additif puis étendue au cas où le bruit serait coloré. L'apport d'une transformation invariante par translation a été observé dans les deux cas précédents.

Pour une évaluation de l'efficacité de cette approche, des tests ont été réalisés sur des images bruitées par un bruit blanc additif ou par un bruit coloré additif. Des mesures sur le gain en rapport signal sur bruit ont été réalisées. Pour compléter notre évaluation des mesures sur le taux de préservation des contours et de la corrélation ont été également effectuées.

On peut améliorer cette méthode en exploitant certaines informations a priori sur les coefficients en ondelettes comme leurs voisinages ou leurs corrélations à travers les échelles.

Bibliographie

- [Bru97] A.G.Bruce et H.Y.Gao. “Waveshrink with firm shrinkage”. *Statistica Sinica*, vol 4, pp 855-874, 1997.
- [Coh90] A.Cohen, I.Daubechies et J.C.Feauveau “Biorthogonal bases of compactly supported wavelets ”, *Commun. Pure. App. Math.*, vol. 45, pp 485-560, 1990
- [Cha95] Y.T.Chan, “Wavelets Basics”, Royal Militairy College of Canada, Edition Kluwer Academic Publishers, 1995.
- [Chu92] C.K.Chui, “An Introduction to Wavelets”, Academic Press, 1992.
- [Cham98] A.Chambolle, R.DeVore, N.Y.Lee et B.J.Lucier, “Nonlinear wavelet image processing : variational problems, compression and noise removal through wavelet shrinkage”, *IEEE transaction on image processing* vol 7, No 3, pp 319-335, 1998.
- [Cro98] M.S.Crouse, R.D.Novak, et R.G.Braniuk, “Wavelet-based signal processing using hidden markov models”, *IEEE transaction on signal processing* vol 64, No 4, special issue on wavelet and filterbanks pp 886-902, 1998.
- [Dau88] I.Daubechies, “ Orthonormal Bases of compactly supported Wavelets ”, *Commun. Pure Appl. Math.*, Vol 46 pp 909–996, 1988
- [Dau94] I.Daubechies, “ Ten Lectures on Wavelets ”, CBMS, SIAM, Vol 61, pp 194–202, 1994.
- [Dja93] J.P. Djamdji “Analyse en ondelettes et mise en correspondance en télédétection ”, Thèse de Doctorat, Université de Nice, Sophia Antipolis, 1994.

- [Don94] D.L. Donoho et I.M. Johnstone “Idéale spatial adaptation via wavelet shrinkage”, *Biometrika*, vol 81, pp 425-455, 1994.
- [Don95a] D.L. Donoho, I.M. Johnstone, G. Kerkycharian et D. Picard “Wavelet shrinkage : asymptotia ?” *Journal of the Royal Statistical Society, Series B*, Vol 57, No 2 pp 301–369, 1995.
- [Don95b] D.L. Donoho, “Denoising by soft–thresholding”. *IEEE Transactions on Information Theory*, vol 41, No 3, pp 613-627, 1995.
- [Don95c] D.L. Donoho et I.M. Johnstone “Adapting to unknown smoothness by wavelet shrinkage”, *J. Amer. Statist. Assoc*, Vol 90, pp : 1200-1224, 1995.
- [Fle81] R. Flecher “Practical methods of optimisation”, Wiley-interscience publications, Wiley, Chichester, 1981.
- [Fal94] F. Falzon, “Analyse multiéchelle, détection des singularité et caractérisation de la régularité des images”, Thèse de Doctorat, université Nice, Sophia Antipolis, 1994.
- [Gab46] D. Gabor “Theory of communication”, *J. Inst. Elect. Eng. (London)* 93 III p429-457, 1946.
- [Gro84] A. Grossmann et J. Morlet “Decomposition of hardy fonctions into square integrable wavelets of constant shape” *SIAM J. Math. Anal.* 15, 723-736, 1984.
- [Jan97] M. Jansen, M. Malfait, et A. Bultheel, “Generalised cross validation for wavelet thresholding”, *Signal Processing*, Vol 56, No1, pp 33-44, 1997.
- [Jan2000] M. Jansen “Wavelet thresholding and noise reduction”, Thèse de Doctorat, Université Catholique de Louvain, Avril 2000.
- [Joh94] I.M. Johnstone et B.W. Silverman “Wavelet threshold estimator for data with correlated noise” Décembre 1994

- [Kun88] M. Kunt, G.Granlund et M.Kocher, “ Traitement Numérique des Images ”, Presse Polytechniques Universitaires Romandes 1988.
- [Mal89] S.G. Mallat “ A theory for multiresolution signal decomposition : the wavelet representation ”, IEEE Trans. Pattn. Anal. Mach. Intell. Vol 11, No7 pp 674-693, 1989.
- [Mey92] Y.Meyer “ Ondelettes : algorithmes et applications ”, Edition Armond Colin 1992.
- [Nas95] G.P.Nason et B.W.Silverman “ The Stationnary Wavelet Transform and some Statistical Applications ”, Rapport Technique, Département de Mathématique, Université de Bristol.1995
- [Nas96] G.P.Nason “ Wavelet shrinkage using cross-validation ”, Journal of Statistical Society, Series B, vol 58, pp 463-479, 1996.
- [She92] M.J.Shensa “ The discrete wavelet transform : wedding the à trous and Mallat algorithms ”, IEEE Trans. Signal Process., vol 40, No 10 , Octobre 1992.
- [Ste81]C. Stein “ Estimation of the mean of a multivariate normal distribution ”, Ann. Stat. , vol 9 , pp 1135-1151, 1981.
- [Wah90] G.Wahba, “Spline models for obsevational data”, CBMS-NSF Regional Conf. Series in Appl Math. Society for Industrial and Applied Mathematics, Philadelphia, PA, 1990.
- [Wey98] N.Weyrich et G.T.Warhola “ Wavelet Shrinkage and Generalized Cross Validation for Image Denoising ”, IEEE, Trans. Image Process. , vol 7, No1, Janvier 1998
- [Xu94] Y.Xu, J.B.Weaver, D.M.Healy et J.Lu, “Wavelet transforme domain filters : a spatially selective noise filtration technique”, IEEE transaction on image processing, vol 3, No 6, pp 747-758, 1994.