ONDELETTES ET ALGORITHMES CONCURRENTS

Yves Meyer

Institut Universitaire de France
CEREMADE, Université Paris-Dauphine
75775 Paris Cedex 16

RÉSUMÉ

Dix ans après leur apparition en traitement du signal et de l'image, les ondelettes ont enfin trouvé la place qui leur revient et l'objet de cette communication est d'indiquer dans quelle situation il convient d'utiliser les "ondelettes temps-échelle" et dans quel cas les algorithmes concurrents, du type "temps-fréquence" sont plus adaptés. Les problèmes qui seront étudiés sont (1) le débruitage des images transmises par le satellite Hubble, (2) celui d'un enregistrement "live" de Brahms et (3) celui de la compression des empreintes digitales (FBI&Scotland Yard).

ABSTRACT

Wavelets are ten years old, as a tool in signal and image processing. Today we are able to delimitate the scope of "time-scale wavelets" and we know that for a large class of problems, "time-frequency wavelets" are better suited. The problems which will be discussed are (1) denoising algorithms applied to extragalactic astronomy images, (2) denoising of a performance of Brahms playing his first Hungarian Dance, recorded on a wax cylinder and (3) compression of fingerprints.

1. Introduction.

Dix ans après que les ondelettes aient été proposées comme un nouvel outil en traitement de l'image et du signal, il convenait de rendre compte au GRETSI des résultats obtenus et des échecs rencontrés. Les deux principales critiques qui étaient adressées aux ondelettes, il y a quelques années, étaient:

-(1) on ne peut rien espérer de bon d'un algorithme à tout faire qui s'applique en aveugle à n'importe quel signal, dans n'importe quelle situation

-(2) les algorithmes basés sur les ondelettes ne bénéficient pas d'une méthodologie statistique suffisante pour pouvoir être appliqués à des situations réelles.

Nous allons voir que la situation aujourd'hui est tout à fait différente car les techniques d'analyse, de débruitage, de compression et de synthèse par algorithmes "temps-échelle" et "temps-fréquence" ont pris en compte les critiques précédentes et ont trouvé leur forme optimale.

Tout d'abord, on ne travaille plus en aveugle mais on utilise des ondelettes adaptées au problème: ce seront soit des ondelettes "temps-échelle", soit des ondelettes "temps-fréquence". Le critère d'adaptation est soit un théorème d'optimalité comme dans l'étude menée par Donoho [1] soit une composante de l'algorithme qui contient un algorithme de décision, comme dans [4].

Les ondelettes temps -échelle conviennent à l'analyse et à la modélisation des signaux présentant de fortes non-stationnarités (la détection des bords d'une image entre dans ce cadre) et l'exemple le plus impressionnant sera celui des multifractales où la non-stationnarité atteint une violence extrême. Les ondelettes temps-fréquence sont particulièrement adaptées aux signaux quasi-stationnaires parole,musique, textures des images et empreintes digitales).

L'introduction d'une méthodologie statistique est due à D.Picard, D.Donoho et I.Johnstone.

L'exposé qui suit comporte cinq parties:

2. Ondelettes (temps-échelle) et multifractales.
3. Ondelettes et débruitage en traitement du signal et de l'image (application au débruitage optimal des images du sattelite Hubble)
4. Bases trigonométriqes locales et traitement des signaux quasi-stationnaires
5. Paquets d'ondelettes et compression des empreintes digitales.
6. Débruitage de l'enregistrement "live" de J.Brahms interprétant un arrangement pour piano de sa "première danse hongroise".

2. Ondelettes et signaux non-stationnaires: les cas des multifractales.

Le problème posé est d'élucider l'éventuelle structure multifractale du signal mesuré par Yves Gagne dans la soufflerie de Modane. Il s'agit de turbulence développée à haut nombre de Reynolds et la mesure est effectuée en un point fixe du tunnel, par la méthode du fil chaud. Le signal mesuré est donc la variation du voltage nécessaire pour maintenir constante la température du fil, alors qu'il est refroidi par l'écoulement turbulent. Le problème posé est de savoir si ce signal est une fonction multifractale du temps. En acceptant l'hypothèse de Taylor, cela impliquerait l'existence des cascades de Richardson. Grâce aux travaux de U.Frisch et G.Parisi, nous disposons d'une définition précise d'un signal multifractal. En tout point τ du signal, on définit l'exposant de Holder ponctuel $\alpha(\tau)$ par lim-sup $\log|f(t)-f(\tau)|/\log|t-\tau|$ quand t tend vers τ. Le fait que cet exposant de Holder ponctuel soit une fonction très irrégulière du temps signifie que le signal présente de très violentes non-stationnarités.

Nous allons quantifier cette information de non-stationnarité par la définition d'une structure multi-fractale. On calcule à cet effet la mesure de Hausdorff $D(\alpha)$ de l'ensemble des τ tels que $\alpha(\tau) = \alpha$ et l'on dit que le signal est multifractal si cette dimension $D(\alpha)$ est une fonction concave de α sur un certain intervalle $[\alpha_0,\alpha_1]$ et si cette fonction concave est la transformée de Legendre de la valeur moyenne de la puissance p-ième des incréments du signal.
Frisch et Parisi proposent un algorithme par ondelettes pour calculer le "spectre de singularités", c'est à dire la fonction $D(\alpha)$. Cet algorithme stabilise, en un certain sens, le calcul direct à partir des incréments, en les remplaçant par une version adoucie de ces mêmes incréments, c'est à dire par la transformée en ondelettes du signal. Les travaux sont conduits par A.Arnéodo et par un groupe de ses élèves, sur le plan de l'analyse du signal de turbulence, et par S.Jaffard en ce qui concerne la validation de la méthode. Il va sans dire qu'aucune technique de traitement du signal ne permettait d'évaluer les exposants fractaux ponctuels et que l'analyse par ondelettes est la seule méthode qui convienne [2], [4].

3. Ondelettes, compression des images et débruitage.

Considérons tout d'abord, en suivant D.Donoho, le problème général suivant: on dispose (dans le cas unidimensionnel) de données de la forme $f(t_i)= A[g(t_i)]+n(t_i)$, $1\leq i\leq N$, A désignant un opérateur linéaire modélisant l'appareil de mesure et $n(t_i)$ désignant un bruit blanc. Il s'agit donc de mesures indirectes, en présence de bruit. On veut retouver $g(t)$, $0\leq t\leq 1$, grâce à une connaissance a priori sur le signal $g(t)$. La discussion se généralise facilement en dimension 2 et 3. Dans les exemples que nous traiterons cette connaissance a priori sera, soit la régularité de la fonction $g(t)$, soit la régularité en dehors d'un ensemble fini de points (non connus à l'avance) où des discontinuités sont envisagées, soit (et cet exemple est imporant en recherche pétrolière) une hypothèse sur la variation totale de la fonction $g(t)$.
L'opérateur A modélise l'appareillage de mesure et introduit un certain lissage. Numériquement cela signifie que la matrice correspondante est mal conditionnée et il n'est pas envisageable de résoudre le problème en inversant A ce qui aurait pour effet d'accentuer le bruit.
Les méthodes habituellement employées dans ce type de problèmes sont des régularisations de Tychonov. On considère l'opérateur A^*A, que l'on diagonalise dans une base orthonormée, notée $\phi_0, \phi_1, \phi_2,\ldots$. Les valeurs propres correspondantes sont appelées ω_k. On résout le problème mal posé $F=Af+\sigma Z$ (f est la fonction inconnue appartenant à un certain ensemble convexe C et Z est un bruit blanc normalisé) en le remplaçant par $A^*F=A^*Af$

et en décomposant la fonction inconnue f dans la base $\phi_0, \phi_1, \ldots$ On est conduit à $f=\Sigma(1/\omega_k)[F,A\phi_k]\phi_k$ mais cette série n'a aucun sens. Pour lui en donner un, il faut introduire une "fenêtre de régularisation", c'est à dire une pondération μ_k sur les vecteurs ϕ_k.

L'estimateur g de f est alors donné par la série $g(t)=\Sigma(\mu_k/\omega_k)[F,A\phi_k]\phi_k(t)$.

La pondération par les μ_k exprime le crédit de confiance que l'on peut apporter à la composante de l'information F dans le "canal fréquentiel" défini par le vecteur ϕ_k. On choisit donc μ_k proche de 1 quand k est petit et μ_k nul quand k est grand. Comme le remarque Donoho, une régularisation de Tychonov revient à adapter, une fois pour toutes, la base orthonormée à l'opérateur (c'est à dire à l'appareillage) sans tenir compte des propriétés spécifiques du signal ou de l'image que l'on analyse. Or pour une certaine classe d'opérateurs et en faisant une hypothèse a priori sur le signal ou l'image (images géométriques dont les propriétés sont bien décrites dans une certaine gamme d'espaces de Besov), Donoho montre que l'on obtient des résultats bien meilleurs en adaptant le mode de description à ce type de signaux ou d'images-c'est à dire en utilisant une description en ondelettes. Pour définir ces notions de qualité de reconstruction et d'optimalité, Donoho introduit le risque minimax défini par $\inf_g \sup_C E[\|g-f\|^2]$ où C désigne la connaissance a priori sur f, où g est l'estimateur de f et où E est l'espérance mathématique relative au bruit aléatoire (la signification de ce minimax est que dans le pire des cas, l'estimateur g doit donner un bon résultat). On recherche des solutions optimisant la vitesse à laquelle ce risque minimax tend vers 0 quand σ tend vers 0. La solution proposée par Donoho et qui améliore la régularisation de Tychonov repose sur l'existence d'une base orthonormée d'ondelettes ψ_λ, $\lambda \in \Lambda$, et d'une suite u_λ de fonctions appelées vaguelettes dans [1],telles que $A^*u_\lambda=k_\lambda\psi_\lambda$. On a alors formellement $f=\Sigma_\Lambda[Af, u_\lambda]k_\lambda^{-1}\psi_\lambda = \Sigma c_\lambda\psi_\lambda$ et il suffirait de remplacer Af par les données observées F pour conclure. Mais là encore la série diverge et il faut introduire une régularisation. Cette régularisation s'appelle "wavelet shrinkage" et elle consiste à remplacer les coefficients d'ondelettes c_λ par $\mathrm{sign}(c_\lambda)(c_\lambda-t_\lambda)_+$ où les t_λ sont calibrés une fois pour toutes en utilisant la connaissance a priori, à l'aide de formules très simples que l'on trouvera dans [1].

Ce que Donoho nous apprend est qu'il faut préférer un algorithme efficace où l'opérateur que l'on doit étudier se présente sous une forme presque-diagonale à un algorithme lent où l'opérateur soit exactement diagonalisé.

Une application du point de vue que nous venons d'exposer concerne le débruitage [réalisé par A.Lannes et S.Roques, Obsevatoire Midi-Pyrénées] des images fournies par le telescope spatial de Hubble (HST). Comme chacun sait, les images du HST sont dégradées par une forte aberration sphérique du miroir primaire, limitant la résolution à laquelle on pouvait s'attendre. La déconvolution des images du HST présente de nombreuses difficultés dues au caractère non-isotrope et variable dans le champ de la réponse impulsionnelle [3].

De plus 12% de l'énergie, au lieu des 60% attendus, se retrouve dans le pic central dont la largeur est environ 0,1 seconde d'arc.

Le débruitage s'obtient en deux étapes. On doit d'abord estimer la fonction de régularisation qui est ici l'équivalent de la suite que nous avons appelée μ_k. Le rapport signal/bruit, exprimé point par point dans l'espace de Fourier, intervient explicitement dans la formulation de la fonction de régularisation. Mais comment distinguer à partir de l'image transmise par le satellite ce qui est image et ce qui est bruit? C'est là que l'analyse par ondelettes intervient: on étudie (en fonction des échelles dyadiques) les variances des différentes "composantes de détails" de l'image donnée. Ces variances commencent par décroître pour ensuite croître et c'est ce qui permet de savoir quand le bruit commence à masquer le signal. Le lecteur est renvoyé à [3] où l'algorithme est décrit en détail.

4. Bases trigonométriques locales et signaux quasi-stationnaires.

Une analyse de Fourier instantanée est la recherche d'un spectre instantané pour un signal donné. Nous allons nous limiter aux signaux quasi-stationnaires et essayer de définir une segmentation optimale en composantes stationnaires. Cette décomposition optimale sera fondée sur un critère entropique.

En d'autres termes, nous allons, à tout instant τ, chercher un intervalle de confiance $[\tau-\alpha,\tau+\alpha]$ sur lequel le signal semble stationnaire et tester cette stationnarité grâce à un développement du signal étudié $s(t)$, sur cet intervalle, en série de Fourier. Le problème est la recherche de ces intervalles de confiance et la définition précise des atomes temps-fréquence utilisés pour reproduire le signal complet grâce à un recollement des structures locales mises en évidence dans l'approche quasi-stationnaire.

Nous allons présenter un algorithme qui prenne en compte le programme précédent mais nous ne savons pas s'il fournit une solution optimale. On commence par rechercher les intervalles $[\tau_j, \tau_{j+1}]$ fournissant la segmentation adaptative du signal quasi-stationnaire en intervalles de confiance. Pour cela, on part d'une partition de la droite réelle délimitée par des nombres réels τ_j , $j=0, \pm 1, \pm 2, \ldots$ On pose $\delta_j=\tau_{j+1}-\tau_j$ et l'on se donne une suite η_j de réels positifs tels que $\eta_j+\eta_{j+1} \leq \delta_j$. On associe à cette partition des fenêtres $w_j(t)$ ayant les propriétés suivantes:

$w_j(t)=1$ sur $[\tau_j+\eta_j, \tau_{j+1}-\eta_{j+1}]$,

$w_j(t)=0$ si $t \leq t_j-\eta_j$ ou si $t \geq \tau_{j+1}+\eta_{j+1}$.

$w_j(\tau_j+s)=w_{j-1}(\tau_j-s)$ si $|s|\leq\eta_j$.

$|w_j(\tau_j+s)|^2+|w_j(\tau_j-s)|^2=1$ si $|s|\leq\eta_j$.

Ces fenêtres serviront à segmenter le signal donné. Les atomes temps-fréquence que nous utiliserons sont alors donnés par $w_{j,k}(t)=(2/\delta_j)^{-1/2}\sin[(\pi(k+1/2)(t-\tau_j)/\delta_j]w_j(t)$, $j= 0, \pm 1, \pm 2, \ldots$, $k=0, 1, 2, \ldots$

Le résultat fondamental, obtenu dans des cas particuliers par H.Malvar, M.Vetterli etc., est que cette suite double $w_{j,k}(t)$ est une base orthonormée de l'espace des signaux d'énergie finie. On dispose alors d'une infinité de bases orthonormées (paramétrée par l'ensemble des choix possibles des fenêtres de segmentation) et le choix de la base optimale se fera sur un critère entropique. Si les coefficients de la décomposition d'un signal [dont l'énergie est normalisée à 1] sont notés $c_{j,k}$, on cherchera la base qui minimise $-\sum\sum|c_{j,k}|^2\log|c_{j,k}|^2$. On aura ainsi comprimé de façon optimale le signal [4]. On peut montrer que le temps de calcul de la base optimale est en NlogN pour un signal échantillonné sur N points. Ces algorithmes se généralisent en dimension 2 ou 3 et l'image est alors segmentée de façon optimale en rectangles. Dans le cas de l'image, la technique précédente est une analyse de textures quasi-stationnaires.

5. Les paquets d'ondelettes.

Tout comme les bases trigonométriques locales constituent une amélioration sensible de la DCT, les paquets d'ondelettes sont une version du système de Walsh, amélioration permettant une meilleure définition fréquentielle. Rappelons la définition dy système de Walsh (utilisé depuis longtemps en traitement du signal). On part de la fonction $r(t)$, périodique de période 1, définie par $r(t)=1$ sur $[0,1/2)$, $r(t)=-1$ sur $[1/2,1]$. Si $n=\varepsilon_0+\varepsilon_1 2+\ldots+\varepsilon_j 2^j$, alors $W_n(t)=[r(t)]^{\varepsilon_0}[r(2t)]^{\varepsilon_1}\ldots[r(2^jt)]^{\varepsilon_j}$ .Naturellemnt cette base orthonormée de $L^2[0,1]$ fournit également la base $W_n(t-k)$, $n=0,1\ldots$, $k=0,+1,+2$ ...de $L^2(\mathbb{R})$.

La généralisation que nous avons en vue repose sur une seconde construction équivalente du système de Walsh. On a

$W_{2n}(t)= W_n(2t)+W_n(2t+1)$ et

$W_{2n+1}(t)= W_n(2t)-W_n(2t+1)$,

$n=0,1,2\ldots$. Si au lieu des coefficients 1,1,1 et -1, on utilise deux filtres miroirs en quadrature conduisant aux ondelettes orthonormées d'Ingrid Daubechies, on généralise le système de Walsh en imposant les conditions

$W_{2n}(t)=h_0W_n(2t)+h_1W_n(2t+1)+\ldots+h_{N-1}W_n(2t-N+1)$ et

$W_{2n+1}(t)=g_0W_n(t)+g_1W_n(2t+1)+\ldots+g_{N-1}W_n(2t-N+1)$.

Ceci définit par récurrence les fonctions $W_n(t)$, à condition d'imposer la condition initiale $\int W_0(t)dt=1$. A partir de ces fonctions $W_n(t)$, appelées paquets d'ondelettes de base, on construit les fonctions $W_{n,j,k}(t)=2^{j/2}W_n(2^jt-k)$, $j=0, \pm 1, \pm 2,\ldots$, $k=0, +1, +2,\ldots$, $n=0,1,2,\ldots$.

Toute partition de $[0,\infty)$ par des intervalles dyadiques $[n2^j, (n+1)2^j)$ conduit à une base orthonormée par les fonctions $W_{n,j,k}(t)$ correspondantes.
On dispose ainsi d'une vaste collection de bases orthonormées et le choix de la meilleure base se fait à l'aide du critère entropique déjà évoqué. Il s'agit donc d'un filtrage adaptatif. La compression par paquets d'ondelettes est le standard adopté par le FBI pour l'archivage des empreintes digitales.

6. Le denoising de l'enregistrement "live" de Brahms.

Les deux méthodes de segmentation que nous avons décrites sont, en quelque sorte duales l'une de l'autre. Dans un cas, il s'agit d'une segmentation adaptative et dans le second cas d'un filtrage adaptatif. Il y a cependant des situations où aucune des deux n'est satisfaisante. On emploie alors la méthode itérative suivante. Considérons un signal $s(t)$, échantillonné sur N points et décomposé dans la base orthonormée $w_1, w_2,\ldots w_N$. On a $s(t)=\sum \alpha_i w_i(t)$ et l'on suppose $\sum|\alpha_i|^2=1$. On définit le <u>facteur de compression</u> par $\delta(s)=d(s)/N$ où $d(s)=2^{\varepsilon(s)}$, $0 \leq d(s) \leq N$, et où $\varepsilon(s)=-\sum|\alpha_i|^2\log|\alpha_i|^2$. L'algorithme fonctionne alors comme suit. On cherche la meilleure base $w_1, w_2,\ldots,w_N$ permettant de décomposer $s(t)$. On définit alors $g_M(t) = \alpha_1 w_1(t)+\ldots+\alpha_M w_M(t)$ et $r_M(t)= s(t)-g_M(t)$. On choisit M comme le plus petit entier tel que la compression de $r_M(t)$, (renormalisé en le divisant par $\|r_M(t)\|$) dans la base choisie dépasse un certain seuil τ_0: $\delta(r_M/\|r_M\|) \geq \tau_0$.
Cela signifie qu'en demeurant dans la base qui représente le mieux le signal, ce reste (une fois normalisé pour ramener son énergie à 1) est perçu comme un bruit. On itère cette procédure sur le reste $r_M(t)$, considéré comme un nouveau signal. On stoppe l'algorithme quand la dimension théorique du n-ième reste dépasse τ_0.
Le grand avantage de cette approche est la possibilité à chaque itération de changer complètement de méthode d'analyse. C'est cette approche au débruitage qui fournit les meilleurs résultats sur l'enregistrement "live" de Brahms réalisé sur un disque de cire fourni par Thomas Edison...Les détails de cette aventure se trouvent dans [5].

1.**D.Donoho**. Wavelet shrinkage and Wavelet-Vaguelette decomposition, Actes du colloque de Toulouse, Y.Meyer & S.Roques ed., Editions Frontières (1993).
2.**S.Jaffard**. Formalisme multifractal pour les fonctions.C.R.Acad.Sc.Paris, à paraître.
3.**A.Lannes et S.Roques**. Preprint. Observatoire Midi-Pyrénées. Toulouse.
4. **Y.Meyer**. Wavelets, algorithms and applications. SIAM (1993).
5.**W.Wayt Gibbs**. Making wavelets. Scientific American, June 1993, p.105.