Régularité ponctuelle, Analyse multifractale, et Traitement des signaux

Jacques Lévy Véhel

▶ **To cite this version:**

Jacques Lévy Véhel. Régularité ponctuelle, Analyse multifractale, et Traitement des signaux. 1997. inria-00593297

HAL Id: inria-00593297
https://inria.hal.science/inria-00593297

Submitted on 13 May 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Régularité Ponctuelle, Analyse Multifractale, et Traitement des signaux

Jacques LÉVY VÉHEL

INRIA
Domaine de Voluceau
Rocquencourt
B.P.105
78153 Le Chesnay Cedex
Jacques.Levy_Vehel@inria.fr
http://www-syntim.inria.fr/fractales/

Table des matières

1 Régularité Ponctuelle et Application en Synthèse de la Parole **3**

1.1 Dimension de boîte locale 3
1.2 Dimension de Hausdorff locale 4
1.3 Exposant de Hölder ponctuel 4
1.4 Exposant de Hölder local 5
1.5 Analyse 2–microlocale 6
1.6 Lien entre exposant de Hölder ponctuel et dimension de boîte locale 7
1.7 Prescription de la régularité Höldérienne ponctuelle 7
1.8 Prescription de la frontière 2–microlocale 8
1.9 Estimation 8
1.9.1 Méthode fondée sur la transformée en ondelettes 9
1.9.2 Utilisation des distributions de Wigner-Ville 10
1.9.3 Estimation de l'exposant de Hölder ponctuel par GIFS 12
1.10 Représentation d'un signal par analyse de la régularité locale 18
1.10.1 Introduction 18
1.10.2 Analyse et synthèse 19
1.10.3 Application à des signaux de parole 21

2 Analyse Multifractale et Traitement d'images **43**

2.1 Spectre de Hausdorff 43
2.2 Spectre de grandes déviations 43
2.3 Spectre de Legendre 45
2.4 Comparaison des spectres 45
2.5 Exemple : le mouvement Brownien fractionnaire (mBf) 46
2.6 Analyse multifractale de mesures 48
2.6.1 Exposant de Hölder 48
2.6.2 Spectre de Hausdorff 48
2.6.3 Spectre de grandes déviations 48
2.6.4 Spectre de Legendre 49
2.6.5 Exemple : la mesure binomiale 49
2.7 Application : Traitement des images 50
2.7.1 Généralités 53
2.7.2 Description multifractale des images 55

Introduction

Le but de ce cours est d'exposer certaines méthodes d'analyse de signaux "irréguliers" et d'en montrer des applications dans les domaines de la parole et de l'image. Les signaux fractals sont des signaux complexes qui possèdent une certaine forme d'ordre ou de structure ; c'est cette structure qui leur confère un rôle particulier parmi l'ensemble des signaux irréguliers : elle permet de dériver des méthodes naturelles pour les analyser, méthodes qui sont souvent fondées sur la transformée en ondelettes. Une fois les outils mis au point, on peut les appliquer à des classes de signaux plus générales, en gardant à l'esprit l'archétype fractal comme guide.

1 Régularité Ponctuelle et Application en Synthèse de la Parole

Il existe de nombreuses façons de mesurer l'irrégularité locale d'un signal. Nous nous intéressons dans la suite à des dimensions locales et aux exposants de Hölder et 2–microlocaux. Pour simplifier les notations, la discussion est conduite pour des fonctions à valeurs dans $\mathbb{R}$.

1.1 Dimension de boîte locale

Soit F le graphe d'une fonction f à valeurs dans $\mathbb{R}$. Soit N_δ^ε le nombre de carrés de la forme $[i\delta,(i+1)\delta[\times[j\delta,(j+1)\delta[, (i,j)\in\mathbb{N}^2, \delta>0$, qui intersectent la partie de F correspondant aux points $y\in]x-\varepsilon, x+\varepsilon[$.

On définit les dimensions de boîte locale inférieures et supérieures de F en x par :

$$\begin{aligned}\overline{\dim}_B^x F &= \lim_{\varepsilon\to 0}\limsup_{\delta\to 0} -\frac{\log N_\delta^\varepsilon}{\log\delta}\\ \underline{\dim}_B^x F &= \lim_{\varepsilon\to 0}\liminf_{\delta\to 0} -\frac{\log N_\delta^\varepsilon}{\log\delta}\end{aligned}$$

Quand ces nombres coïncident, on parle de dimension de boîte locale de F en x.

Exemple : soit $f(x) : x^\alpha \cos\frac{1}{x^\beta}, x\neq 0, f(0)=0$ avec $0<\alpha<\beta$.

Alors : $\dim_B^x = 2-\frac{\alpha+1}{\beta+1}$

1.2 Dimension de Hausdorff locale

Soit encore F le graphe d'une fonction f à valeur dans $\mathbb{R}$. Pour $\delta > 0, r \geq 0$, on définit

$$H^r_\delta(F) = \inf \left\{ \sum_{i=1}^{\infty} |E_i|^r, \{E_i\}_{i\in\mathbb{N}} \delta\text{-recouvrement de } F \right\}$$

où un δ-recouvrement est une collection dénombrable de boules de diamètre inférieur ou égal à δ recouvrant F.

H^r_δ est une fonction décroissante de δ, et on note :

$$H^r(F) = \lim_{\delta\to 0} H^r_\delta(F)$$

H^r est la mesure de Hausdorff de F en dimension r. On montre qu'il existe un unique réel $\dim_H F$, la dimension de Hausdorff de F, tel que :

$$\begin{aligned} r < \dim_H F &\implies H^r(F) = +\infty \\ r > \dim_H F &\implies H^r(F) = 0 \end{aligned}$$

Pour définir la dimension locale, on pose :

$$\dim^x_H(F) = \lim_{\varepsilon\to 0} \dim_H F^x_\varepsilon$$

où F^x_ε est la restriction du graphe de f correspondant aux points d'abscisse appartenant à $B(x,\varepsilon) =]x-\varepsilon, x+\varepsilon[$.

1.3 Exposant de Hölder ponctuel

L'idée est ici de définir un exposant α tel que f se "comporte" comme $|x-x_0|^\alpha$ au voisinage de x_0. C'est une sorte de généralisation de la dérivée ; quand f est C^1, on sait que $f(x) - f(x_0)$ est, au premier ordre, proportionnel à $x - x_0$. Quand f n'est pas dérivable, on cherche si on peut écrire $|f(x) - f(x_0)| \sim c|x-x_0|^\alpha$. Plus généralement, on regarde si :

$$c_1|x-x_0|^\alpha \leq |f(x) - f(x_0)| \leq c_2|x-x_0|^\alpha \text{ pour } 0 < c_1 \leq c_2$$

ce qui permet par exemple de rendre compte du cas $f(t) = |t|^\alpha(\sin t + 2)$.

Cependant, une telle condition est trop restrictive et ne permet pas de définir α dans toutes les situations. Pour traiter des cas comme $f(t) : |t|^\alpha \log|t|$, il faut écrire :

$$c_1|x-x_0|^{\alpha+\varepsilon} \leq |f(x) - f(x_0)| \leq c_2|x-x_0|^{\alpha-\varepsilon}$$

Enfin, il existe des situations où la première inégalité ne sera pas vérifiée en tout x : par exemple $f(t) = |t|^\alpha \sin \frac{1}{t}$. On impose donc cette condition seulement pour une suite :

$$\exists \{x_n\}, x_n \to x_0 / c_1 |x_n - x_0|^{\alpha+\varepsilon} \leq |f(x_n) - f(x_0)|$$

L'étape suivante est de considérer les cas où f est dérivable, mais où une dérivée d'ordre supérieure de f n'existe pas : dans ce cas, il faut retrancher à f son polynôme de Taylor pour voir apparaître la singularité et se retrouver dans le cas précédent.

Finalement, on définit α, l'exposant ponctuel de Hölder de f en x_0 par :

- $\forall \gamma < \alpha, \lim_{h \to 0} \frac{|f(x_0+h) - P(h)|}{|h|^\gamma} = 0,$
- si $\alpha < +\infty, \forall \gamma > \alpha, \limsup_{h \to 0} \frac{|f(x_0+h) - P(h)|}{|h|^\gamma} = +\infty$

où P est un polynôme de degré inférieur ou égal à la partie entière de α. On dit alors que $f \in C^\alpha_{x_0}$.

Cette définition n'est valable que pour $\alpha \in \overline{\mathbb{R}}^{+*} \setminus \mathbb{N}$. Si α est entier, elle doit être adaptée.

Une définition équivalente est, quand $0 < \alpha < 1$:

$$\exists c, \rho_0 > 0, \forall \rho < \rho_0 \quad \sup_{x,y \in B(x_0,\rho)} |f(x) - f(y)| \leq c\, \rho^\alpha$$

Exemple : soit $f(t) = |t|^\alpha \sin \frac{1}{t^\beta}, f(0) = 0, \qquad 0 < \alpha < \min(1, \beta).$

Alors : $\alpha(0) = \alpha$.

1.4 Exposant de Hölder local

L'exemple précédent montre que l'exposant de Hölder ponctuel, comme son nom l'indique, mesure la régularité ponctuelle de f en un point donné x_0. Dans certains cas, il est désirable de tenir compte du comportement de f dans un *voisinage* de x_0, et d'incorporer cette information dans l'exposant.

On définit alors un exposant local $\widetilde{\alpha}$ comme suit :

$$\exists c, \rho_0 > 0, \forall \rho < \rho_0 \quad \sup_{x,y \in B(x_0,\rho)} \frac{|f(x) - f(y)|}{|x - y|^{\widetilde{\alpha}}} \leq c$$

($\widetilde{\alpha}$ est bien sûr le plus grand réel tel que l'inégalité ci-dessus est vérifié.) Cette définition est valable pour f non dérivable, sinon il faut comme d'habitude retrancher le polynôme de Taylor.

Exemple : reprenons $f(t) = |t|^\alpha \sin \frac{1}{t^\beta}, f(0) = 0, \qquad 0 < \alpha < \min(1, \beta)$
alors : $\widetilde{\alpha}\,(0) = \frac{\alpha}{1+\beta}$.
Sur cet exemple, nous voyons que :

$$\widetilde{\alpha}\,(0) = \frac{\alpha}{1+\beta} < \alpha(0) = \alpha < 2 - \dim_B^0(f) = \frac{\alpha+1}{\beta+1} \text{ si } \alpha\beta < 1$$

Remarque 1

On peut en fait définir une infinité d'exposants dans le même esprit que α et $\widetilde{\alpha}$. C'est l'objet de l'analyse 2–microlocale.

1.5 Analyse 2–microlocale

L'analyse 2–microlocale fournit une description plus fine de la régularité ponctuelle/locale que le seul exposant de Hölder. Elle est fondée sur une analyse de Littlewood–Paley, mais nous exposons ici une définition fondée sur les coefficients d'ondelettes continus $T(a,b)$ ou discrets s_j^k.

Définition 1

$$\begin{aligned} & f \in C_{x_0}^{s,s'} \ ssi \\ & \forall 0 < a + |b - x_0| < 1 \\ & |T(a,b)| \leq c\ a^{s+s'}(a + |b - x_0|)^{-s'} \end{aligned}$$

ou :

$$|s_j^k| \leq c\ 2^{-js}(1 + |k - 2^j x_0|)^{-s'}$$

Les propriétés essentielles des espaces 2–microlocaux concernent leur stabilité sous l'action d'opérateurs différentiels, et leur rapport avec les espaces de Hölder :

Proposition 1

$$f \in C_{x_0}^{s,s'} \Longleftrightarrow f' \in C_{x_0}^{s-1,s'}$$

$$\begin{aligned} s + s' > 0 \implies & C_{x_0}^{s,s'} \subset C_{x_0}^{s} \subset C_{x_0}^{s,-s} \\ & C_{x_0}^{s,s'} \subset C_{x_0}^{s-\varepsilon,s'+\varepsilon}\ \forall\ \varepsilon > 0 \end{aligned}$$

Cette dernière relation, couplée au fait que le domaine 2–microlocal d'une fonction (i.e. les couples (s,s') tels que $f \in C_{x_0}^{s,s'}$) est un sous-ensemble convexe de $\mathbb{R}^2$, implique que la totalité de ce domaine est déterminée par sa frontière. Cette frontière sera exprimée dans la suite par la paramétrisation suivante : $s = s(\sigma)$ où $\sigma = s + s'$.

1.6 Lien entre exposant de Hölder ponctuel et dimension de boîte locale

Nous avons vu, sur l'exemple du chirp, un cas où $\alpha(x) < 2 - \dim_B^x$. D'un autre côté, si $f(x) = x$, alors $\alpha(x) = +\infty > 2 - \dim_B^x = 1$. Enfin, $\alpha(0) = 2 - \dim_B^0 F$ pour $|t|^\alpha \sin \frac{1}{t^{1/\alpha}}, 0 < \alpha < 1$.

Dans quels cas peut-on comparer $\alpha < 1$ et $\dim_B^x$?

Proposition 2 *Soit*

$$\begin{aligned} \overline{c}(t,\varepsilon) &= \inf\{c \in \mathbb{R}^{+*} / \forall u \in B(t,\varepsilon), |f(t) - f(u)| < c|t-u|^{\alpha(t)}\} \\ \underline{c}(t,\varepsilon) &= \sup\{c \in \mathbb{R}^{+*} / \exists u \in B(t,\varepsilon), |f(t) - f(u)| \geq c\, \varepsilon^{\alpha(t)}\} \end{aligned}$$

Si :

$$\exists \varepsilon' > 0 / \sup_{t \in B(x,\varepsilon)} \overline{c}(t,\varepsilon) < +\infty \qquad \forall \varepsilon < \varepsilon' \tag{1}$$

Alors :

$$\overline{\dim_B^x} F \leq 2 - \min\{\liminf_{t \to x} \alpha(t), \alpha(x)\}$$

Si

$$\exists \varepsilon' > 0 / \inf_{t \in B(x,\varepsilon)} \underline{c}(t,\varepsilon)(t,\varepsilon) > 0 \qquad \forall \varepsilon < \varepsilon' \tag{2}$$

Alors :

$$\underline{\dim_B^x} F \geq 2 - \max\{\limsup_{t \to x} \alpha(t), \alpha(x)\}.$$

En particulier, si 1 et 2 sont vraies, et si la fonction $x \to \alpha(x)$ est continue en x, alors $\dim_B^x F$ existe et vaut $2 - \alpha(x)$.

Ceci nous amène à étudier la fonction $x \to \alpha(x)$, que nous appelons fonction de Hölder de f.

1.7 Prescription de la régularité Höldérienne ponctuelle

On se pose le problème suivant :

Soit s une fonction de $[0,1]$ dans $[0,1]$. Sous quelles conditions sur s existe-t-il une fonction f continue de $[0,1]$ dans $\mathbb{R}$ telle que $\alpha_f(x) = s(x)$ pour tout x ?

Une des raisons pour résoudre ce problème et construire des fonctions à régularité prescrite (c'est-à-dire trouver f connaissant s) est que, dans certaines applications comme la

synthèse de la parole (voir plus bas), il est désirable de contrôler précisément la régularité ponctuelle du signal construit, parce que celle-ci contient une information essentielle, indépendante par exemple de l'amplitude du signal. Le résultat principal est le suivant :

Théorème 1 *Soit s une fonction de $[0,1]$ dans $[0,1]$. Les conditions suivantes sont équivalentes :*

- *s est la fonction de Hölder d'une fonction continue.*
- *il existe une suite de fonctions continues $(s_n)_{n\geq 1}$, telles que :*

$$s(x) = \liminf_{n\to\infty} s_n(x)$$

Il existe plusieurs méthodes pour construire f à partir de s admissible. Trois de ces méthodes sont exposées en détail dans l'article joint.

1.8 Prescription de la frontière 2–microlocale

Pour contrôler plus précisément le comportement local de f, on peut chercher à imposer non seulement l'exposant de Hölder de f en x_0, mais son appartenance aux ensembles 2–microlocaux $C^{s,s'}_{x_0}$.

On peut alors montrer que :

Proposition 3 *Une fonction $s(\sigma)$ est la frontière 2–microlocale d'une fonction f si et seulement si $s(\sigma)$ est concave et décroissante. Dans ce cas, la fonction f définie par :*

$$f = \sum_{j>0}\sum_{k} c_{j,k}\Psi_{j,k}$$

avec

$$c_{j,k} = \inf_{\sigma}(2^{-j\sigma}(2^{-j} + |k\ 2^{-j}|)^{s(\sigma)-\sigma})$$

admet $s(\sigma)$ comme frontière 2–microlocale en $x = 0$.

1.9 Estimation

Si l'on veut utiliser l'information fournie par la régularité Höldérienne dans des applications pratiques en traitement du signal, il est essentiel de pouvoir estimer α à partir de données discrètes. Disons tout de suite qu'il n'existe pas, à l'heure actuelle, de méthode fiable qui le permette en tout généralité, même pour des signaux non bruités.

Cela tient à plusieurs facteurs :

- le principal est le caractère discret des données. Comme l'optimalité de $\alpha(x)$ dépend d'une suite $x_n \to x$, on peut très bien se retrouver dans des situations où tous les points d'echantillonnages sont "loins" des x_n, et donc font croire à une régularité plus élevée.

- La fonction de Hölder peut être, on l'a vu au paragraphe précédent, extrêmement irrégulière, par exemple partout discontinue. Il est clair que, sans hypothèses supplémentaires sur $\alpha(x)$, toute méthode d'estimation rencontrera de sérieuses difficultés.

Pour l'instant, deux cas peuvent être traités avec un minimum de succès : celui des singularités isolées, en utilisant un estimateur à base d'ondelettes, et celui où $\alpha(x)$ est continue, en utilisant une modélisation par GIFS. Cependant, dans les deux cas, les résultats ne sont satisfaisants que si $\alpha = \widetilde{\alpha}$ (on dit parfois que la singularité est non oscillante).

En présence de singularités oscillantes ou dans le cas de signaux aléatoires, des estimateurs ont été proposés récemment dans un cadre paramétrique (mouvement Brownien multifractionnaire par exemple), mais nous ne les aborderons pas ici.

1.9.1 Méthode fondée sur la transformée en ondelettes

La méthode repose sur le théorème suivant :

Théorème 2
Soit $\psi(t)$ une fonction lipschitzienne de la variable réelle, vérifiant :

- $|\psi(t)| \leq C(1+|t|)^{-3}$; $|\psi'(t)| \leq C(1+|t|)^{-3}$
- *la collection $\psi_{(j,k)}(t) = 2^{j/2}\psi(2^j t - k)$, où j et k décrivent l'ensemble des entiers, est une base orthonormée de $L^2(\mathbb{R})$.*

Soit f une fonction telle que :

$$\int_{-\infty}^{+\infty} |f(t)|(1+|t|)^{-3}dt < +\infty\,,$$

les coefficients d'ondelettes de f sont définis par :

$$c(j,k) = \int_{-\infty}^{+\infty} f(t)\psi_{(j,k)}(t)dt.$$

On a alors les deux résultats suivants :

- *si en un point t_0, f vérifie :*

$$|f(t) - f(t_0)| \leq C|t - t_0|^s$$

pour un certain $s \in]0\,;1]$, alors :

$$|c(j,k)| \leq C' 2^{-j(s+1/2)}(1 + |2^j t_0 - k|^s) \tag{3}$$

pour tout j et k.

- *Si (3) est vérifiée et s'il existe $\epsilon > 0$ tel que $f \in C^\epsilon$, alors, pour tout t tel que $|t - t_0| \leq 1$, on a :*

$$|f(t) - f(t_0)| \leq C|t - t_0|^s \log \frac{2}{|t - t_0|}$$

Ce résultat montre que l'exposant de Hölder d'une fonction f, soumise à une condition de régularité globale, peut s'évaluer, à une correction logarithmique près, à l'aide des modules des coefficients d'ondelette de f, via une régression dans un diagramme bi-logarithmique.

1.9.2 Utilisation des distributions de Wigner-Ville

Les distributions d'énergie temps-échelle représentent un champ d'étude général, dans lequel le scalogramme (module carré de la transformée en ondelettes) est un cas particulier. Elles peuvent être utilisées pour l'estimation de l'exposant de Hölder, où elles donnent une plus grande liberté de choix des paramètres de contrôle. Une définition standard est fondée sur la distribution de Wigner-Ville et s'écrit :

$$\Omega_f(t, a\pi) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} W_f(u,v)\pi\left(\frac{u-t}{a}, a\ v\right) du\ dv$$

où

$$W_f(t,v) = \int_{-\infty}^{+\infty} f(t + \tau/2) f^*(t - \tau/2) e^{-2\pi i v\tau} d\tau$$

désigne la transformée de Wigner-Ville, et $\pi(t,v)$ est une fonction de paramétrisation arbitraire.

Le scalogramme est un cas particulier de cette classe de distributions. En effet, lorsqu'on choisit $\pi(t,v) = W_\psi(t,v)$, où ψ est une ondelette, on retrouve :

$$\Omega_f(a,t,\pi) = |T_\psi f(a,t)|^2.$$

Il existe dans cette classe d'autres estimateurs non biaisés, c'est-à-dire vérifiant :

$$\mathrm{E}(\Omega_f(t,a,\pi)) \sim a^{2H(t)+1} \qquad \text{quand} \qquad a \to 0$$

où $H(t)$ est l'exposant de Hölder au point t, sous la condition suffisante :

$$\tilde{\pi}(\xi, \xi/2) = 0 \tag{4}$$

où $\tilde{\pi}$ est la transformée de Fourier partielle par rapport au temps de la fonction π. Pour des fonctions π satisfaisant (4) on a :

$$H(t) = \frac{1}{2} \lim_{a \to 0^+} \frac{\log E(\Omega_f(t,a,\pi))}{\log a} - \frac{1}{2}$$

Le cas particulier des fonctions π du type $\pi(t,v) = g(t)H(v)$ est d'un intérêt majeur car il permet de contrôler indépendamment les comportements temporels et fréquentiels (dans ce cas g joue le rôle d'un filtre passe-bas et H celui d'un filtre passe-bande). La condition (4) se traduit alors par :

$$G(\xi).H(\xi/2) = 0$$

où G est la transformée de Fourier de g. De plus la fonction régularisante g peut être choisie de sorte à offrir une décroissance rapide de la variance sur l'estimation des exposants d'échelle sans sacrifier le biais. En particulier, en considérant l'ondelette de Morlet h_M et en choisissant g dans la classe des gaussiennes, la fonction de paramétrisation associée peut s'écrire :

$$\pi(t,v) = (g * W_{h_M}(.,v))(t,v)$$

et la distribution associée prend la forme :

$$\Omega_f(t,a,\pi) = |Tf(.,a)|^2 * \left(\frac{1}{a} g\left(\frac{\cdot}{a}\right)\right)(t,a)$$

qui peut être considérée comme un lissage dépendant de l'échelle du scalogramme. Dans ce cas, chaque échantillonnage d'une distribution donnée, à une échelle fixée, peut être vu comme une moyenne en temps local du scalogramme autour du point t. Par suite, sous des hypothèses de variation lente de $H(t)$ on peut considérer $\Omega_f(a,t,\pi)$ comme un estimateur de la moyenne d'ensemble du scalogramme :

$$\Omega_f(a,t,\pi) = \hat{E}(|Tf(a,t)|^2)$$

Ainsi le problème d'estimation des exposants peut être étudié de manière efficace en utilisant les distributions d'énergie temps-échelle.

1.9.3 Estimation de l'exposant de Hölder ponctuel par GIFS

Rappels sur les GIFS

Nous commençons par rappeler quelques notations utilisées dans l'article joint.
Considérons la collection des ensembles $(F^k)_{k\in\mathbf{N}^*}$, où F^k est un ensemble fini non vide de contractions S_i^k sur $[0\,;1]$, pour $i=0,\ldots,N_k-1$, où $N_k\geq 1$ est un entier qui désigne le cardinal de F^k. On note c_i^k le facteur de contraction de S_i^k, pour tout $i=0,\ldots,N_k-1$, et $k\in\mathbf{N}^*$. Considérons l'opérateur W^k, défini sur tout compact $A\subset[0\,;1]$, par :

$$W^k(A)=\bigcup_{i=0}^{N_k-1}S_i^k(A).$$

Alors, sous certaines conditions, il existe un unique compact G tel que :

$$\lim_{k\to\infty}W^k\circ\ldots\circ W^1(A)=G\ \ \forall\ A\subset[0\,;1].$$

On appelle G l'attracteur du GIFS $(K,\{F^k\}_{k\in\mathbf{N}^*})$ et $W^k\circ\ldots\circ W^1(A)$ son attracteur à l'itération k (qui dépend de A).
Dans la suite nous considérons le cas où les S_i^k sont des contractions affines et le nombres des contractions à l'itération k, soit N_k, est égal à m^k, pour un certain entier $m\geq 2$. Nous supposons plus précisément que les transformations $S_i^k\,(0\leq i<m^k)$ peuvent être représentées sous une forme matricielle comme suit :

$$S_i^k\begin{pmatrix}t\\x\end{pmatrix}=\begin{pmatrix}1/m & 0\\ a_i^k & c_i^k\end{pmatrix}\begin{pmatrix}t\\x\end{pmatrix}+\begin{pmatrix}i/m^k\\ b_i^k\end{pmatrix}$$

et que chaque S_i^k opère de $\left[[\frac{i}{m}]m^{-k+1}\,;([\frac{i}{m}]+1)m^{-k+1}\right]$ sur $[im^{-k}\,;(i+1)m^{-k}]$.
Nous noterons :

$$S_i^k(t,x)=(L_i^k(t),F_i^k(t,x))$$

où

$$\begin{cases}L_i^k(t)=\frac{i}{m}t+\frac{i}{m^k}\\[2ex] F_i^k(t,x)=a_i^kt+c_i^kx+b_i^k\end{cases}$$

Principe de la méthode

Par souci de simplicité, nous nous restreignons ici à des fonctions définies et continues sur $[0\,;1]$. L'algorithme de calcul de l'exposant de Hölder est fondée sur la proposition suivante, démontrée dans l'article joint, que nous rappelons ici :

Proposition 4 *Sous certaines conditions, l'attracteur du GIFS défini ci-dessus est le graphe d'une fonction continue f telle que :*

$$f\left(\frac{i}{m}\right) = y_i \quad \forall\, i = 0, \ldots, m$$

et

$$\alpha_f(t) = \min(\alpha_1, \alpha_2, \alpha_3)$$

où

$$\begin{cases} \alpha_1 = \liminf\limits_{k\to+\infty} \frac{\log(c^k_{m^{k-1}i_1+m^{k-2}i_2+\ldots+mi_{k-1}+i_k} \ldots c^2_{mi_1+i_2} c^1_{i_1})}{\log(m^{-k})} \\ \alpha_2 = \liminf\limits_{k\to+\infty} \frac{\log(c^k_{m^{k-1}j_1+m^{k-2}j_2+\ldots+mj_{k-1}+j_k} \ldots c^2_{mj_1+j_2} c^1_{j_1})}{\log(m^{-k})} \\ \alpha_3 = \liminf\limits_{k\to+\infty} \frac{\log(c^k_{m^{k-1}l_1+m^{k-2}l_2+\ldots+ml_{k-1}+l_k} \ldots c^2_{ml_1+l_2} c^1_{l_1})}{\log(m^{-k})} \end{cases} \tag{5}$$

et où les i_p, j_p et les l_p's sont définis comme dans la proposition ci-dessus.

Pour appliquer cette proposition au calcul de l'exposant de Hölder d'un signal réel continu f, il faut d'abord calculer les coefficients c^j_k du GIFS dont l'attracteur est f. Ceci revient à résoudre le "problème inverse" pour les GIFS, qui est une généralisation du classique problème inverse pour les IFS. Ce dernier problème a été étudié par de nombreux auteurs, soit d'un point de vue théorique, soit dans des applications en physique, en compression d'images, ou en traitement du signal. Le problème inverse s'énonce comme suit :

"Etant donné un signal f, trouver un (G)IFS dont l'attrateur approxime au mieux f, au sens d'une norme fixée."

Ce problème, extrêmement complexe dans le cas des IFS, est beaucoup plus simple à résoudre pour les GIFS. En particulier, la proposition 4 implique que, pour une fonction $f \in C^0([0\,;1])$, on peut trouver un GIFS dont l'attracteur est aussi proche que l'on veut, au sens de la norme $\|.\|_\infty$, de f. Ceci n'est évidement pas le cas quand on se restreint à des IFS dont le nombre de fonctions est fini. Le prix à payer pour une telle simplification et une telle amélioration de l'approximation est évidemment le passage d'un nombre fini à un nombre infini de paramètres dans la modélisation.
La résolution pratique du problème inverse pour les GIFS est exposée dans la prochaine section. Remarquons pour l'instant que, dans le cas particulier où

$$\begin{cases} m = 2 \\ \\ L^k_i(t) = \frac{i}{m} t + \frac{i}{m^k} \end{cases}$$

si l'on pose :

$$s_k^j = f\left((k+\frac{1}{2})2^{-j}\right) - \frac{1}{2}\left(f(k2^{-j}) + f((k+1)2^{-j})\right) \ \forall k \in \{0,...,2^j-1\},$$

et

$$c_k^j = \frac{s_k^j}{s_{[\frac{k}{2}]}^{j-1}} \ \forall k = 0,...,2^j-1, \tag{6}$$

alors, comme m et les L_i^k sont fixés, le problème inverse est résolu si les c_k^j vérifient les conditions (c) et (c'). En effet, dans ce cas, l'attracteur du GIFS défini par les points d'interpolation $\{(0,0)\,;(\frac{1}{2},f(\frac{1}{2}))\,;(1,0)\}$ et les coefficients c_i^j est le graphe de f. Pour calculer la fonction de Hölder de f, on note :

$$\mathcal{S} = \left\{ f \in C^0([0\,;1]) \ : |c_k^j| \in [\frac{1}{2}+\epsilon\,;1-\epsilon] \right\}$$

où les c_k^j sont déterminés par (6) et ϵ est un réel strictement positif fixé (aussi petit que l'on veut). Si $f \in \mathcal{S}$, nous sommes dans les conditions d'application de la proposition 4 et les exposants de Hölder ponctuels de $f \in \mathcal{S}$ sont obtenus en appliquant la relation (5).

Algorithme

Soit $f \in C^0([0\,;1])$. Supposons pour simplifier les notations que $f(0) = f(1) = 0$. f s'écrit alors :

$$f(x) = \sum_{j\geq 0} \sum_{0\leq k<2^j} s_k^j \theta_{j,k}(x).$$

La méthode décrite dans la section précédente ne permet d'obtenir l'exposant de Hölder de f que si celle-ci appartient à $\mathcal{S}$. Dans le cas général, on ne peut espérer calculer α_f mais seulement une approximation. Pour ce faire, on construit une fonction $\tilde{f}$ vérifiant :

- $\tilde{f} \in \mathcal{S}$
- $\forall x \in [0\,;1], \forall g \in \mathcal{S} \ : |\alpha_f(x) - \alpha_{\tilde{f}}(x)| \leq |\alpha_f(x) - \alpha_g(x)|$

et on calculera $\alpha_{\tilde{f}}$ au lieu de α_f.
Cette démarche ressemble d'une certaine manière à celle qui consiste à filtrer un signal avant de l'échantillonner : si la fréquence d'échantillonnage ω_e est inférieure à $2\omega_{max}$, où ω_{max} désigne la bande passante du signal, alors l'échantillonnage direct du signal conduira à une détérioration de l'information. Tout ce que l'on peut récupérer quand $\omega_e < 2\omega_{max}$ est la partie "basse fréquence" du signal, et la bonne manière pour ce faire est de considérer

une approximation du signal définie dans le domaine fréquentiel avant d'effectuer l'échantillonnage. Dans notre cas, nous recherchons une approximation $\tilde{f}$ de f dans le sens décrit ci-dessus, et nous calculons ensuite $\alpha_{\tilde{f}}$. Bien évidemment, cette approche ne donnera de bons résultats que si f est bien approximable selon le critère que nous nous sommes défini. Nous verrons dans les applications que c'est assez souvent le cas en pratique. Insistons sur le fait que $\tilde{f}$ est en général une mauvaise approximation de f au sens de la norme $\|.\|_{L^2}$ ou $\|.\|_\infty$, et que c'est seulement $\alpha_{\tilde{f}}$ qui est proche de α_f.
On vérifie aisément que, sous les hypothèses ci-dessus, $\tilde{f}$ est donnée comme l'attracteur du GIFS défini par :

- $m = 2$
- $L_i^k(t) = \frac{i}{m}t + \frac{i}{m^k}$
-

$$\tilde{c}_k^j = \begin{cases} c_k^j & \text{si} \quad |c_k^j| \in]\frac{1}{2}\,;1[\\ 1-\epsilon & \text{si} \quad |c_k^j| > 1-\epsilon \\ \frac{1}{2}+\epsilon & \text{si} \quad |c_k^j| < \frac{1}{2}+\epsilon \end{cases}$$

Application

Soit $F = \{f(i), i = 0, ..., 2^J\}$ un signal discret donné. Pour $j = 1, ..., J$, considérons l'ensemble P_j défini par :

$$P_j = \{f(i2^{J-j}), i = 0, ..., 2^j\}.$$

P_j n'est autre que le signal sous-échantillonné avec le pas 2^{J-j}. Une interprétation géométrique du calcul des c_k^j est la suivante : Pour $j \in \{1, ..., J-1\}$, l'ensemble $\{c_k^j, k = 0, ..., 2^j - 1\}$ des coefficients qui vont déterminer le GIFS est obtenu comme étant l'ensemble des pentes des 2^j fonctions affines qui permettent de transformer le polygone défini par P_j en celui défini par P_{j+1}. Pour clarifier, nous montrons dans la figure 1 un exemple avec $J = 3$. Les échantillons du signal sont représentés par des petits cercles. Pour $j = 1, 2$, les coefficients du GIFS sont :

$$c_k^j = \frac{u_k^j}{u_{[\frac{k}{2}]}^{j-1}} \; \forall k = 0, ..., 2^j - 1.$$

La procédure d'estimation de l'exposant de Hölder ponctuel est la suivante :

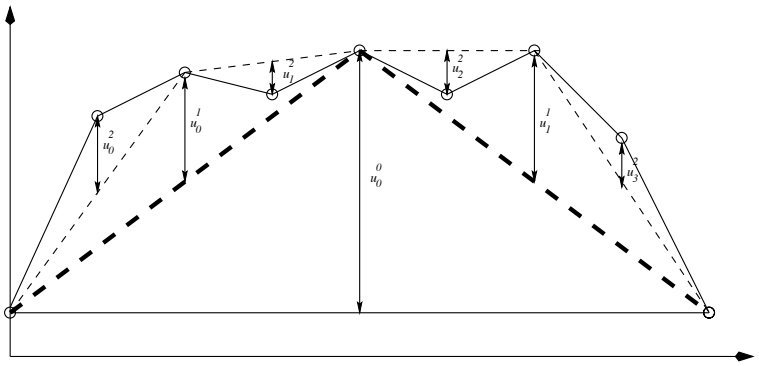

FIG. 1 – *Exemple de calcul des* c_k^j.

Soient $i \in \{0, ..., 2^J - 1\}$, et $(i_1, ..., i_{J-1})$ l'uplet tel que $i2^{-J} = \sum_{p=1}^{J-1} i_p 2^{-p}$. Posons $k_j = \sum_{p=0}^{j-1} i_{j-p} 2^p$ et considérons la séquence $(C^j)_{j=1,...,J-1}$ definie par :

$$C^j = -\log_2(|c_{k_j}^j| \dots |c_{k_2}^2||c_{k_1}^1|).$$

Nous définissons :

$$\tilde{C}^j = -\log_2(|\tilde{c}_{k_j}^j| \dots |\tilde{c}_{k_2}^2||\tilde{c}_{k_1}^1|).$$

L'exposant $\alpha_f(i)$ est alors donné comme le résultat de la regresion linéaire de $\tilde{C}^j$ en fonction de j.

Remarquons que si $\tilde{c}_k^j = c_k^j$, pour tout k et j, i.e tous les coefficients du GIFS ont leur module dans $]\frac{1}{2}\,;1[$, alors nous nous retrouvons dans le cadre de l'estimation de l'exposant de Hölder à partir des coefficients de Schauder s_k^j. En effet, dans ce cas $\tilde{C}^j = -\log_2|s_{k_j}^j|$ et il est connu que la base de Schauder caractérise la régularité locale des fonctions continues nulle part dérivables (et même de fonctions plus régulières, sous certaines conditions).

Nous décrivons maintenant des tests numériques sur des signaux synthétiques. Nous comparons les trois méthodes décrites ci-dessus (Ondelettes, Wigner-Ville et GIFS) sur des fonctions de Weierstrass généralisées :

$$W(t) = \sum_{k=0}^{\infty} 2^{-ks(t)} sin(2^k t),$$

pour lesquelles la fonction de Hölder est $\alpha_W(t) = s(t)$, si s est continue. Le choix de ces fonctions est motivé par plusieurs raisons :

– elles sont simples à synthétiser et il est facile de faire varier $\alpha_W(t)$,

– elles ont déjà été considéré par plusieurs auteurs et fournissent un bon modèle générique pour différents phénomènes,

– elles ne sont pas synthétisées à partir d'ondelettes ou de GIFS, ce qui permet de tester de façon indépendante les algorithmes d'estimation.

La figure 2 (resp. 4) montre la fonction $W(t)$ sur $[0\,;1]$ obtenue en prenant $s(t) = t$ (resp. $s(t) = |sin(5\pi t)|$). La figure 3 (resp. 5) montre l'estimation de sa fonction de Hölder utilisant les différentes méthodes citées précédement, la courbe théorique étant en noir. La courbe en jaune représente l'estimation en utilisant la distribution d'énérgie temps-échelle (pseudo-Wigner affine). La courbe en bleu représente l'estimation en utilisant le scalogramme, i.e une transformée en ondelettes continue. L'ondelette utilisée est l'ondelette de Morlet. La courbe en vert représente l'estimation en utilisant la base de Schauder. Finalement, la courbe en rouge représente l'estimation par GIFS.

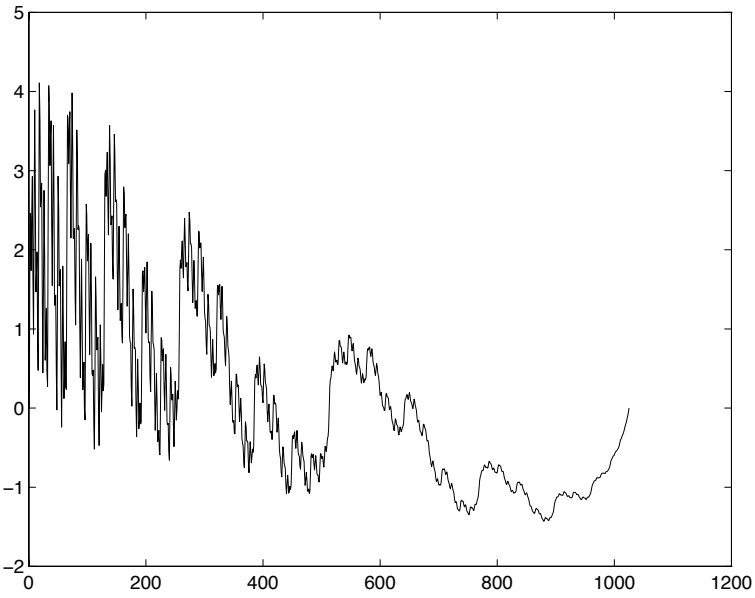

FIG. 2 – *Fonction de Weierstrass généralisée avec* $s(t) = t$.

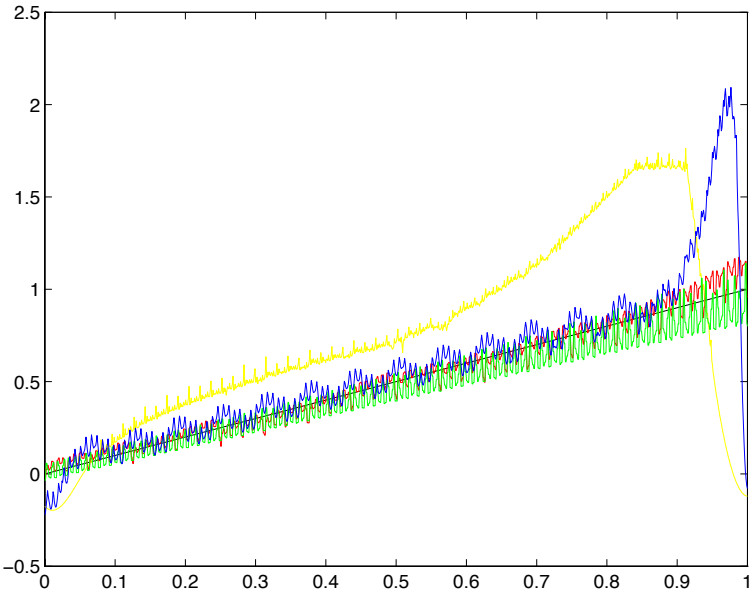

FIG. 3 – *Estimation de la fonction de Hölder.*

1.10 Représentation d'un signal par analyse de la régularité locale

1.10.1 Introduction

Alors que la littérature abonde en méthodes d'estimation de l'exposant de Hölder, il existe peu de travaux qui utilisent directement l'information donnée par les exposants de Hölder d'un signal pour bâtir un modèle de ce signal. En d'autres termes, l'information concernant la régularité locale est utilisée pour l'interprétation et l'analyse des signaux mais pas pour leur synthèse.

Dans cette section nous allons développer une technique qui permet de donner une représentation fonctionnelle d'un signal donné, à partir de sa fonction de Hölder. Cette démarche peut paraître surprenante *a priori* puisque la connaissance de la fonction de Hölder d'un signal ne permet pas de caractériser ce signal : au contraire, l'ensemble des signaux ayant une fonction de Hölder fixée est dense dans $C^0([0\,;1])$ (voir article joint). Ceci a deux conséquences : nous devons nous appuyer sur une information additionnelle pour reconstruire le signal et nous ne cherchons pas à obtenir une synthèse exacte, mais seulement une approximation plus ou moins précise suivant la conformité du signal à notre modèle.

La méthode repose une fois encore sur la modélisation d'un signal par GIFS. Soit $(c_k^j)_{k,j}$ les coefficients du GIFS qui permet de reconstruire le signal. On calcule des approximations α_j

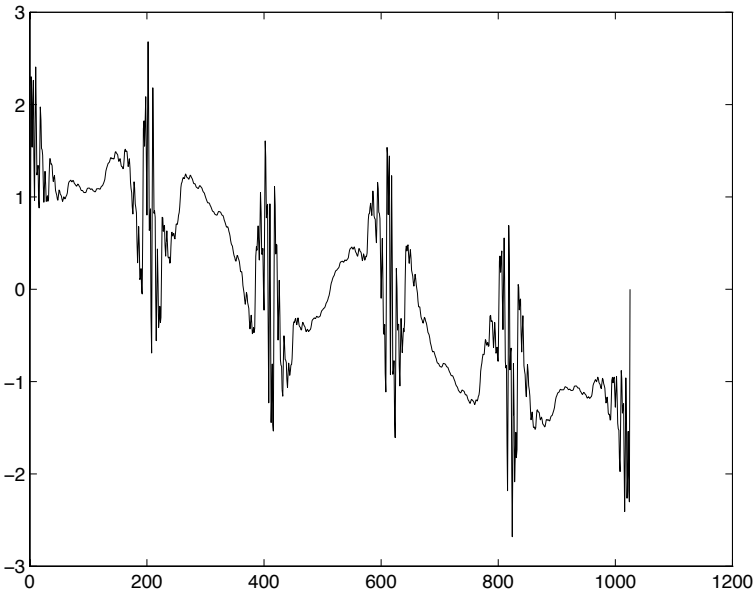

FIG. 4 – *Fonction de Weierstrass généralisée avec* $s(t) = |sin(5\pi t)|$.

à différentes résolutions j (i.e en ne tenant en compte que des c_k^l pour $l \leq j$) de la fonction de Hölder du signal. On obtient alors une nouvelle représentation, équivalente à celle donnée par les c_k^j. Schématiquement, si le signal original f contient 2^J échantillons, on peut écrire :

$$\begin{aligned} \left\{f(k2^{-J}), k = 0, ..., 2^J - 1\right\} &\iff \left\{c_k^j \,; j = 1, ..., J-1 \,; k = 0, ..., 2^j - 1\right\} \\ &\iff \overbrace{\left\{c_k^j \,; j = 1, ..., J - (j_0 + 1) \,; k = 0, ..., 2^j - 1\right\}}^{C} \\ &\quad \bigcup \qquad \underbrace{\{\alpha_j \,; j = J - j_0, ..., J-1\}}_{A} \end{aligned}$$

A ce stade, nous savons donc reconstruire le signal à partir de points d'interpolation, définis par C, et de régularités locales à différentes échelles, décrites dans A. Pour que cette representation soit plus qu'un artifice de notation et qu'on puisse réellement parler de synthèse à partir de la régularité locale, il nous faut examiner de plus près la structure de A. C'est l'objet de la prochaine section.

1.10.2 Analyse et synthèse

Soit $s(t)$ une fonction *continue* de $[0\,;1]$ dans $]0\,;1[$. Considérons le GIFS dont les coefficients c_k^n sont définis par :

$$c_k^n = 2^{-s(k2^{-n})} \ \forall n \geq 1 \ \forall k \in \{0, ..., 2^n - 1\}.$$

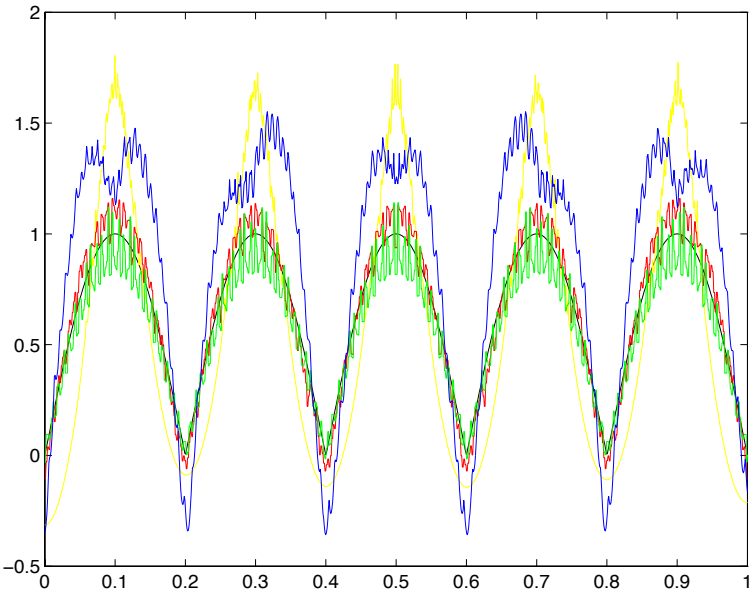

FIG. 5 – *Estimation de la fonction de Hölder.*

Pour tout $t \in [0\,;1]$ et $n \geq 1$, définissons $k_n = k_n(t)$ par :

$$|k_n 2^{-n} - t| = \inf_k |k2^{-n} - t|.$$

Si f est la fonction dont le graphe est l'attracteur de ce GIFS,

$$\alpha_f(t) = \lim_n \frac{s(k_1 2^{-1}) + ... + s(k_n 2^{-n})}{n} = \lim_n s(k_n 2^{-n}) = s(t).$$

Ceci implique particulièrement que, si on désigne par $s_j(t) = s(k_j 2^{-j})$ la fonction de Hölder de f "à la résolution j", alors pour j assez grand, on a : $s_j(t) \sim s_{j+1}(t)$.
Pour revenir au thème évoqué dans la section précédente, en identifiant les s_j et les α_j, nous voyons que l'ensemble A aura une structure simple en particulier si :

- $\alpha_f(t)$ continue,
- $|\alpha_f(k2^{-j}) - \alpha_j(k2^{-j})|$ est "petit" pour tout k et j.

Pour satisfaire au mieux ces 2 conditions, on effectue un filtrage de f qui consiste à lisser tous les α_j de A. Plus précisément, on effectue sur chaque α_j une convolution par un filtre passe-bas qui fournit un $\tilde{\alpha}_j$:

- continu,

– synthétisable par un petit nombre de paramètres, par exemple en utilisant une modélisation par spline.

La synthèse se fait comme suit : nous gardons tous les coefficients du GIFS appartenant à C et, pour tout $j = j_0, ..., 1$, nous posons :

$$\tilde{c}_k^{J-j} = 2^{-\tilde{\alpha}_j\left(k2^{-(J-j)}\right)} \; \forall k = 0, ..., 2^{J-j} - 1$$

En résumé, le signal synthétique est construit à partir d'un sous échantillonnage régulier du signal avec un pas de 2^{j_0}, et d'un nombre petit de paramètres définissant les $\tilde{\alpha}_j$. Plus j_0 est grand, plus la synthèse est économique. Le choix de j_0 se fait selon la qualité souhaitée de l'approximation.

1.10.3 Application à des signaux de parole

Nous développons dans cette section un modèle pour les signaux de paroles fondé sur l'étude des régularités locales. Nous commençons par une présentation générale des signaux de parole et de leur méthodes classiques de traitement.

Genéralités sur la parole et son traitement automatique

Ce paragraphe donne une brève description du signal vocal, ainsi que les grandes lignes de quelques importantes méthodes d'analyse, de synthèse et de reconnaissance de la parole.

Les sons de la parole

La parole est produite soit par les vibrations des cordes vocales, ce qui donne lieu aux sons voisés, soit par une turbulence créée par l'écoulement de l'air dans une constriction du conduit vocal ou lors du relâchement d'une occlusion de ce conduit, ce qui donne lieu aux sons non voisés. Les plus brèves unités sonores qui permettent de distinguer différents mots s'appellent "les phonèmes".
Les voyelles sont classées selon quatre critères :
- la nasalité ,
- le degré d'ouverture du conduit vocal,
- la position de la constriction principale du conduit vocal,
- la protrusion des lèvres.

Les consonnes sont classées selon :
- le voisement ,

- le mode articulatoire (occlusif, nasal, fricatif, glissant ou liquide),
- la position de la constriction principale.

Tous les aspects de la parole n'ont pas pour fonction d'établir des distinctions de signification. Ainsi les phénomènes "prosodiques" servent souvent à exprimer l'attitude du locuteur envers son exposé. Le timbre de la voix, quant à lui, joue un rôle dans l'identification du locuteur. Au niveau acoustique, ces différentes fonctions de la parole sont imbriquées, le même paramètre pouvant servir différentes fonctions suivant le contexte ou l'échelle de temps considérée.

Description du signal de parole

Le signal de parole est un signal réel continu, d'énergie finie, non stationnaire, très évolutif dans le temps, et dont la structure est très complexe. L'intensité du son émis est liée à la pression de l'air en amont du larynx, sa hauteur est fixée par la fréquence F_0 des vibrations des cordes vocales, appelée fréquence du fondamental ou pitch. Elle peut varier de :
- 80 à 200 Hz pour une voix d'homme,
- 150 à 450 Hz pour une voix de femme,
- 200 à 600 Hz pour une voix d'enfant.

Dans le spectre d'un son voisé, on observe des raies qui correspondent aux harmoniques du fondamental F_0. L'enveloppe de ces raies présente des maxima, appelés formants, qui correspondent aux fréquences propres $F_i (i = 1, 2, 3, \ldots)$ du conduit vocal. Les trois premiers formants sont essentiels pour caractériser le spectre, les autres sont moins importants.
Un son non voisé ne présente pas de structure périodique, il peut être considéré comme un bruit blanc filtré par la transmittance de la partie du conduit vocal située entre la constriction et les lèvres. Son spectre ne présente pas donc de structure de pitch.
Deux sons de même intensité et de même hauteur se distinguent par le timbre. Le mot "timbre" peut avoir plusieurs significations : par exemple, le timbre des voyelles qui qualifie la structure physique du phonème, et le timbre du locuteur qui qualifie l'aspect global de la voix. Les amplitudes relatives des harmoniques du fondamental semblent être des paramètres pertinents pour la détermination du timbre d'un locuteur.

Description acoustique de la parole

Classiquement, la description acoustique des sons de parole s'appuie sur la représentation *fréquence-temps* fournie par le spectrographe. L'intensité des composants spectraux est donnée par le degré de noirceur du tracé. Tous les spectrographes classiques permettent aussi la réalisation de représentations *fréquence-intensité* un instant donné, ce qu'on appelle "sections". Mais, depuis plusieurs années, l'analyse numérique par FFT, LPC ou Cepstre ont remplacé la réalisation des "sections" spectrographiques. Nous donnerons les grandes lignes de ces techniques d'analyse dans les paragraphes suivants.

• Les voyelles en français :

La structure acoustique des voyelles en français se caractérise principalement par la présence de maxima spectraux, c'est-à-dire de zones de fréquences où les harmoniques sont particulièrement intenses (les formants). Le modèle de production classique confirme que la fonction de transfert du conduit vocal des voyelles non-nasales ne comporte que des pôles. Les voyelles d'un locuteur masculin possèdent en général 5 formants. Le premier formant est le pic spectral ayant la fréquence la plus basse, le deuxième formant est le pic suivant, et ainsi de suite. On ne tient pas compte du pic de très basse fréquence (formant glottal), vers $200 - 300Hz$, qui peut apparaître pour les voyelles d'intensité faible, comme $/a/$ par exemple. Des expériences de synthèse ont montré que la position fréquentielle des 3 premiers formants caractérisait le timbre vocalique.

• Les consonnes en français :

On peut décomposer les consonnes, au plan acoustique, en trois ensembles ayant des caractéristiques distinctes. Les occlusives, comme $/t/$ $/d/$, se caractérisent principalement par un silence provenant de la fermeture complète du conduit vocal (occlusion) en un lieu bien défini. Les fricatives (ou constrictives), comme $/s/$ et $/z/$, sont des bruits de friction aérodynamique créés par un écoulement d'air turbulent en un point du conduit vocal présentant un fort rétrécissement. Les sonnantes, comme $/m$ et $/j/$, possèdent, comme les voyelles, une structure formantique. Mais au contact des consonnes sourdes (ou non voisées), les sonnantes perdent leur voisement. Dans ce cas, elles apparaissent comme des fricatives dont le bruit serait concentré aux fréquences formantiques. La présence d'une constriction pour les sonnantes implique que l'un des trois premiers formants au moins aura une intensité très faible. Les occlusives et les fricatives du français sont sourdes ou sonores. Ce voisement est phonologiquement distinctif en français, ce qui n'est pas le cas pour les sonnantes.

Méthodes d'analyse spectrale de la parole

La transformée de Fourier rapide (FFT)

La FFT joue un rôle très important dans le traitement de la parole car elle permet d'obtenir des spectres en temps réel. Des analyses successives sur une durée limitée du signal s'obtiennent en prélevant les échantillons de parole à l'aide d'une fenêtre temporelle glissante. En général, les fenêtres successives se recouvrent. Ces fenêtres doivent avoir une longueur suffisante si l'on veut que la FFT ait un sens (en général on prend 256 ou 512 points). Le recouvrement est par exemple de moitié. Ce procédé s'appelle "analyse de Fourier à fenêtres glissantes".

Le cepstre

En utilisant la FFT pour le calcul du spectre, l'intermodulation source/conduit rend difficile la mesure des formants F_i et de la fondamentale F_0 caractéristiques précisément du conduit et de la source. Le lissage cepstral ou *cepstre* est une méthode qui vise à séparer leur contribution respective par déconvolution. Pour cela, on fait l'hypothèse que le signal vocal s_n est produit par un signal d'excitation g_n (source glottique) traversant un système linéaire passif de réponse impulsionnelle b_n (conduit oral et nasal). Avec cette hypothèse, on peut écrire :

$$s_n = g_n * b_n.$$

Pour déconvoluer s_n (c'est-à-dire retrouver g_n et b_n), il faut se donner une classe de fonctions admissible pour g_n : On suppose que g_n est une séquence d'impulsions (périodique pour les sons voisées). Ces hypothèses sont très limitatives, mais dans la pratique cette méthode fournit des résultats acceptables.

Pour déconvoluer plus aisément s_n, il suffit de transposer le problème, par homomorphisme dans un espace où l'opérateur * de convolution correspond à un opérateur +. Soit D_*^+ cet homomorphisme :

$$s_n = g_n * b_n \ \overset{D_*^+}{\rightarrow} \ s_n^* = g_n^* + b_n^*$$

Après séparation de g_n^* et b_n^*, si la transformation inverse est D_+^*, on aura :

$$\begin{cases} \overset{D_+^*}{g_n^* \ \rightarrow \ g_n} \\ \\ \overset{D_+^*}{b_n^* \ \rightarrow \ b_n} \end{cases}$$

L'intérêt de la méthode pour la parole réside dans le fait que g_n^* et b_n^* sont facilement séparable par un filtre temporel et ce grâce à l'hypothèse simplificatrice sur g_n. D_*^+ et son inverse D_+^* peuvent s'exprimer comme suit :

$$D_*^+ = Z(.) \circ \log(.) \circ Z^{-1}(.)$$

$$D_+^* = Z(.) \circ \exp(.) \circ Z^{-1}(.),$$

où Z est la transformée en z. On considère alors l'homomorphisme DM_*^+ et son inverse DM_+^* définis par :

$$DM_*^+ = Z(.) \circ \log |.| \circ Z^{-1}(.)$$

$$DM_+^* = Z(.) \circ \exp |.| \circ Z^{-1}(.).$$

Les coefficients spectraux sont alors définis par :

$$c_n = DM_*^+(s_n).$$

En pratique, on peut remplacer $Z(.)$ par la FFT, ce qui définit de manière analogue les coefficients cepstraux approchés $c_n^\dagger$, qui prennent leurs valeurs dans un domaine pseudo-temporel appelé domaine quéfrentiel, avec :

$$c_n^\dagger = g_n^\dagger + b_n^\dagger$$

où $g_n^\dagger$ et $b_n^\dagger$ sont respectivement les transposées dans le domaine quéfrentiel de g_n et de b_n. La structure de la parole et les hypothèses faites sur g_n permettent de poser que :

- g_n se réduit à une séquence d'impulsions séparées de n_0 échantillons (n_0 correspond à la fréquence fondamentale F_0).
- $b_n^\dagger$ décroît rapidement (en $1/n$) avec n et devient rapidement négligeable, du moins lorsque $n \geq n_0$.

Dans ces conditions, on peut admettre que les premiers coefficients cepstraux contiennent essentiellement la contribution du conduit, et que les "pics" périodiques visibles sur la suite c_n, $n \geq n_0$ reflètent les impulsions de la source.

Les bancs de filtres : le vocodeur à canaux

L'estimation de l'enveloppe spectrale du signal de parole peut se faire à l'aide de filtres en découpant la bande utile en sous-bandes dans lesquelles on évalue l'intensité du signal. A la limite, lorsque les sous-bandes sont de largeur nulle on tend vers une transformée de Fourier si les filtres utilisés sont idéaux. Dans une certaine mesure, les modèles de l'oreille peuvent être rangés dans la catégorie des vocodeurs à canaux, ce qui donne à ceux-ci un grand intérêt. Chaque canal est constitué d'un filtre passe-bande et d'un système de détection de l'enveloppe spectrale. Le calcul de l'enveloppe ou énergie se fait par sommations des valeurs absolues (ou des carrés) du signal filtré sur une fenêtre donnée. L'estimation du spectre est fournie par la suite des valeurs sortant à un instant donné des N canaux du vocodeur.
Le choix des caractéristiques des filtres du vocodeur est un problème capital. De plus en plus, ces filtres ont des caractéristiques se rapprochant de celles de l'audition humaine. Ils donnent des indices spectraux plus robustes que d'autre méthodes car ils sont moins sensibles à la non-synchronisation de la fenêtre d'analyse sur la période fondamentale. Mais leur inertie (temps de montée, temps d'amortissement) peut être un handicap pour la détection des phénomènes brefs.
En synthèse il faut disposer d'une source et la placer en amont du vocodeur. Le gain des filtres est commandé par les paramètres de synthèse et bien que la phase ne soit pas théoriquement perçue, il faut veiller à ce qu'elle soit sensiblement linéaire sur toute la bande utile. Pour ne pas distordre le spectre synthétisé il faut également veiller à ce que chaque canal ne contienne pas plus de deux harmoniques de la source. Cette contrainte n'est pas facile à respecter du fait de la variabilité du fondamental.

Synthèse de parole

Les techniques de synthèse de parole peuvent être classées en trois catégories :
- synthèse directe ;
- synthèse à travers un modèle ;
- simulation du conduit vocal.

Synthèse directe

Ces techniques produisent la parole par mesure directe de l'onde temporelle. Un exemple est la concaténation de formes d'ondes : à partir d'éléments temporels stockés, le processus de synthèse se contente de concaténer un certain nombre de périodes fondamentales pour reconstituer un certain nombre de phonèmes pour former un mot.

La qualité du signal synthétisé par cette technique peut être acceptable pour certaines applications. Toutefois, cette technique ne permet pas de prendre en compte de façon simple les phénomènes d'influence d'un phonème sur l'autre, qui sont essentiels, notamment pour la compréhension des consonnes.
Ceci dit, en améliorant cette technique de telle sorte à contourner les problèmes qu'elle impose, on peut avoir une très bonne qualité de synthèse. Ceci est le cas du système de synthèse PSOLA développé au CNET LANNION. Le principe de système consiste à superposer et additionner des blocs de signal décalés dans le temps selon un décalage variables synchrone des périodes fondamentales. Ceci permet de modifier la fréquence fondamentale et la durée tout en conservant la qualité du timbre original.

Synthèse à travers un modèle

Ces techniques produisent la parole à travers un modèle de production. Les paramètres de ce modèle sont généralement calculés par minimisation d'un critère d'erreur entre le signal original et le signal prédit par ce modèle. La synthèse à prédiction linéaire et la synthèse par formants font parties de cette catégorie. Nous donnerons ici le principe chacune de ces techniques.

Synthèse à prédiction linéaire

Le modèle de simulation du conduit vocal par prédiction linéaire décrit un échantillon de parole comme étant une combinaison linéaire des échantillons précédents. L'algorithme calcule les coefficients de la combinaison de façon à minimiser l'erreur quadratique moyenne entre le signal original et le signal prédit sur une fenêtre donnée. Ces coefficients sont réactualisés régulièrement toute les 5 à 20 ms. De 10 à 15 coefficients sont nécessaires pour atteindre une synthèse de qualité acceptable. Les avantages de cette technique sont nombreux : analyse automatique, intégration facile des algorithmes, fidélité au timbre original. Cependant, ce modèle donne une représentation trop simpliste de la source d'excitation, ce qui donne un caractère "bruyant" à la parole synthétique. De plus, le modèle se prête mal à la représentation de certains sons tels que les nasales et les constrictives voisées, car les fonctions de transfert obtenues à l'analyse ne contiennent que des pôles.

Synthèse par formants

Dans la synthèse par formants, le signal est engendré par un système qui simule la transmittance proprement dire du conduit vocal, définie par ses fréquences de résonance (les

formants). Pour les sons voisés, ce système est excité par une onde périodique. Pour les sons non voisés l'excitation est un bruit blanc.
Chaque formant, caractérisé par sa fréquence de résonance F_k et sa bande passante B_k, est réalisé par un résonateur du second ordre. Trois ou quatre formants au plus sont représentés. Une résonance supplémentaire peut simuler l'effet des formants d'ordres supérieurs. Les résonateurs peuvent être montés en série ou en parallèle.

Simulation du conduit vocal

Avec cette technique, il s'agit de simuler le fonctionnement physique du système de production de parole. Un modèle articulatoire reconstitue en premier lieu la forme du conduit vocal en fonction de la position des organes phonatoires. Le signal vocal est ensuite calculé à l'aide d'une simulation mathématique de l'écoulement de l'air dans le conduit ainsi délimité. La complexité des calculs que demande cette technique conduit soit à des procédés très coûteux et difficile à réaliser en temps réel, soit à des simplifications dégradant la qualité de la synthèse.

Modélisation fractale

La plupart des approches pour la modélisation des signaux de parole utilisent des représentations sinusoïdales pour les parties voisées du signal, et elles considèrent les parties non voisées comme du bruit blanc filtré par la transmission du conduit vocal. Comme les parties non voisées résultent d'un *écoulement turbulent* de l'air dans une constriction du conduit vocal, il peut paraître intéressant d'étudier une approche qui consiste à donner une représentation fonctionnelle fractale de ces parties du signal, fondée sur l'analyse des régularités locales de la forme d'onde. Cette approche est d'autre part motivée par le fait que les singularités locales, mesurées en termes d'exposants de Hölder, contiennent des informations cruciales sur le signal de parole. Notre paradigme est que la plus grande partie de l'information pertinente peut être obtenue à partir d'un "petit" sous-ensemble des échantillons dans la représentation temporelle (points d'interpolation), et d'une approximation grossière de la régularité locale en chaque point d'échantillonnage.

Cette analyse locale des singularités permet en particulier de décrire quelques caractéristiques du locuteur, de construire un modèle parcimonieux du signal de parole, et de le synthétiser avec une bonne qualité. La tâche principale consiste à construire un modèle permettant d'identifier les points d'interpolation, d'estimer la régularité locale, et finalement

de construire une fonction passant par les points d'interpolation et dont les régularité locale est prescrite. Il est évident que deux représentations différentes sont nécessaires, une pour les parties voisées et une pour les parties non voisées (voir figure 6). Les parties voisées des signaux sont lisses exceptés en quelques points où la variation est brusque. Un simple modèle C^∞ par morceaux fournit une bonne représentation dans ce cas. Les parties non voisées sont irrégulières à toutes les échelles, et il est préferable d'utiliser une représentation qui permet de rendre compte de la singularité locale en tout point. Un modèle à base de GIFS est bien adapté à une telle représentation.
Les données ont été fournies par le CNET dans le cadre d'une étude sur la transformation automatique du timbre de la voix. Ces données sont adaptées à la synthèse par concaténation d'unités acoustiques, développée par le CNET dans la méthode PSOLA, et qui utilise un ensemble d'unités de paroles élémentaires stockées dans un dictionnaire : ces unités sont des diphones (un diphone s'étend de la zone stable d'un phonème à la zone stable du phonème suivant). Elles sont enregistrées par un locuteur réel sous la forme de *logatomes*, mots artificiels composés d'un diphone entouré d'un contexte phonétique neutre. En français, on dénombre 37 phonèmes, et un dictionnaire complet est donc composé de 1369 logatomes. Une fois les logatomes segmentés, la méthode de synthèse consiste à superposer et additionner des blocs de signal, décalés dans le temps de façon variable et synchrone avec les périodes fondamentales. On peut ainsi modifier la fréquence fondamentale et la durée tout en conservant la qualité du timbre original.

Analyse

Pour chaque logatome, nous disposons des marques de "pitch" (fréquence de vibration des cordes vocales) et de transitions voisée/non voisée. Dans un premier temps, les parties voisées sont décomposées en plusieurs unités dont les extrémités sont définies par les marques de pitch. Ces unités sont appelées trames voisées. Pour les parties non voisées, les extrémités des unités sont placées à des intervalle fixés (10ms), et ces unités sont appelées trames non voisées.
Il faut ensuite estimer la fonction de Hölder du logatome. Pour les parties voisées, on détermine les points (isolés) de variation brusque de chaque trame voisée, ainsi que les exposants de Hölder en ces points par la méthode de la transformée en ondelettes. Les points qui ne sont pas à variation brusque sont considérés comme des points où le signal est C^∞. Pour les parties non voisées, on utilise l'algorithme d'estimation fondé sur les GIFS et on applique la pocédure décrite ci-dessus.

Synthèse

L'algorithme de synthèse se caractérise par les étapes suivantes :

- Pour les trames non voisées, les points d'interpolation sont choisis à intervalle fixe ($0.25ms$). Les coefficients de contraction du GIFS sont :

$$c_i^k = m^{-k\tilde{s}(im^{-k})}$$

où $m+1$ est le nombre des points d'interpolation. L'attracteur de l'IFS defini ci-dessus est le graphe de la fonction continue f vérifiant $\alpha_f(t) = \tilde{s}(t)$.

- Chaque trame voisée est modélisée par la fonction :

$$f(t) = \sum_{i=1}^{N} a_i |t - x_i|^{s_i}$$

où les (x_i, y_i) sont les coordonnées des points à variation brusque, les s_i sont leurs exposants de Hölder, et N leur cardinal. Les a_i sont la solution du système linéaire :

$$f(x_i) = y_i \quad i = 1, \ldots, N.$$

La fonction est C^∞ en tout point, sauf aux points x_i, où $\alpha_f(x_i) = s_i$, pour tout $i = 1, \ldots, N$.

Résultats expérimentaux

Les figures 6 et 8 représentent le logatome original /SEUTEUTEU/ prononcé par deux locuteurs mâles, respectivement "ob" et "rg". Les rectangles délimitent les parties voisées. Les figures 7 et 9 montrent respectivement les fonctions de Hölder $\alpha_f(t)$ des locuteurs "ob" et "rg". Par souci de clarté, les valeurs des exposants de Hölder supérieures à 2 sont tronquées à 2. Remarquons que dans la première partie non voisée, $\alpha_f(t)$ varie autour de 0.1 pour le locuteur "ob", et autour de 0.5 pour le locuteur "rg". Pour ces parties, on voit que $\alpha_f(t)$, aux points où le signal n'est pas C^∞, prend ses valeurs dans $[0\,;1]$ pour le locuteur "ob", et dans $[\frac{1}{2}\,;1.8]$ pour le locuteur "rg". On en conclut que la forme d'onde du locuteur "rg" est plus régulière que celle du locuteur "ob". Ceci est en accord avec les tests auditifs où la voix du locuteur "rg" a toujours été qualifiée de plus "douce" que celle du locuteur "ob". La figure 10 montre la fonction lissée $\tilde{s}(t)$ de la première partie non voisée du locuteur "ob". La figure 11 représente la reconstruction du /SEUTEUTEU/ prononcé par le locuteur "ob" pendant 1376 ms, à une fréquence de $16KHz$, ce qui donne un ensemble de 22016

échantillons. Pour les parties voisées, nous avons détecté 4538 points à variations brusques. En prenant en compte les exposants de Hölder, 9076 points doivent être stockés em mémoire. Pour les parties non voisées, un quart des échantillons est stocké em mémoire. Le nombre de paramètres définissant la fonction lisée $\tilde{s}(t)$ pour les quatre parties non voisées est négligeable (10 paramètres).

Ainsi, un ensemble de données d'environ $\frac{N}{3}$ points ($N = 22016$) est suffisant pour reproduire la forme d'onde synthétique. Cette parole synthétique est perceptuellement indiscernable du signal de parole original. La valeur du rapport signal sur bruit entre le signal original et le signal recontruit est de l'ordre de 20 DB.

Les figures 12 et 13 montrent respectivement un zoom du premier /TEU/ du locuteur "ob" et sa reconstruction. On voit que la forme d'onde reconstruite est très proche de l'original.

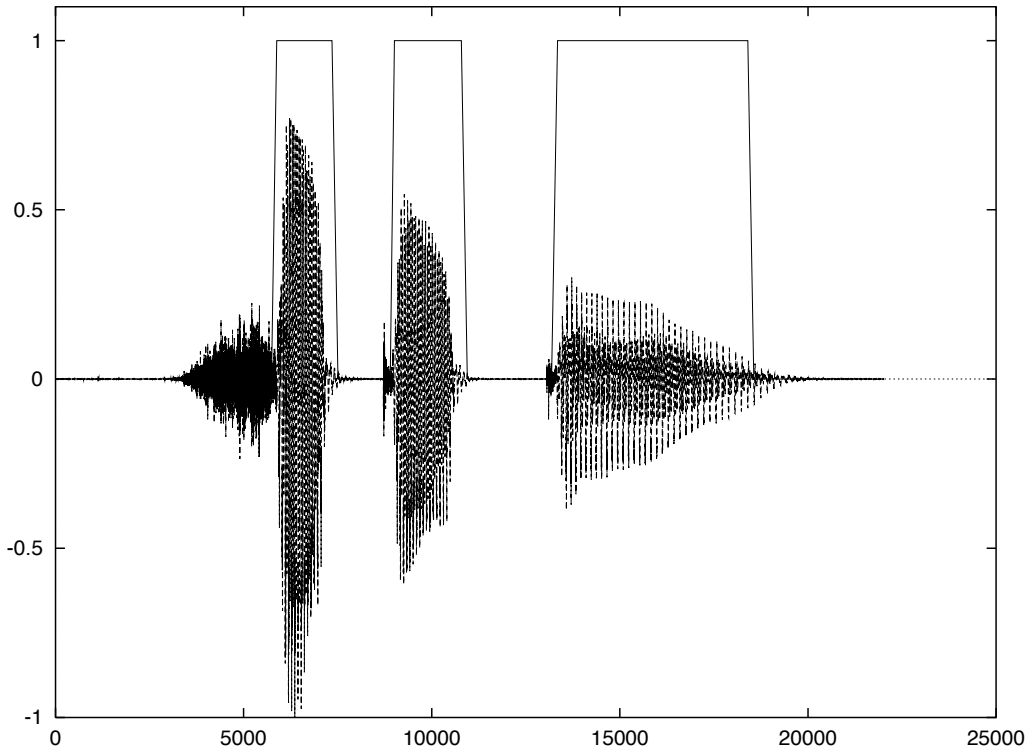

FIG. 6 – */SEUTEUTEU/ (locuteur "ob").*

Application à un signal résiduel de parole obtenue à partir du modèle HNS

Nous présentons dans ce paragraphe une approche qui utilise une représentation différente du signal, fondée sur un modèle, appelé HNS, qui effectue une décomposition harmonique

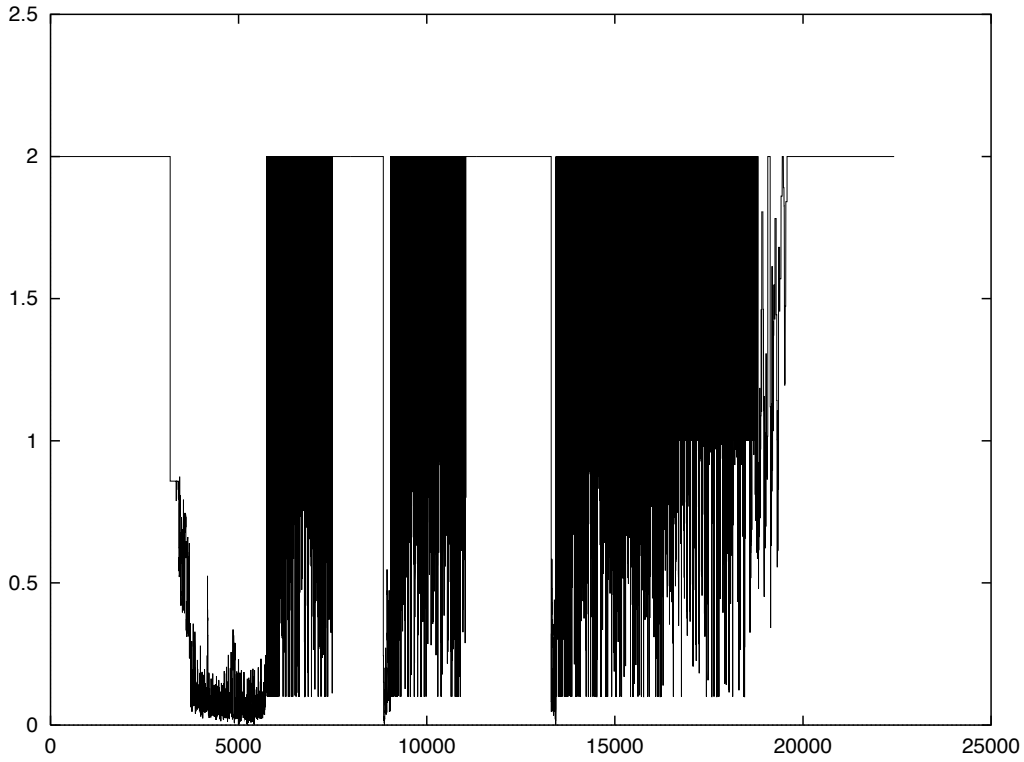

FIG. 7 – *Fonction de Hölder du locuteur "ob".*

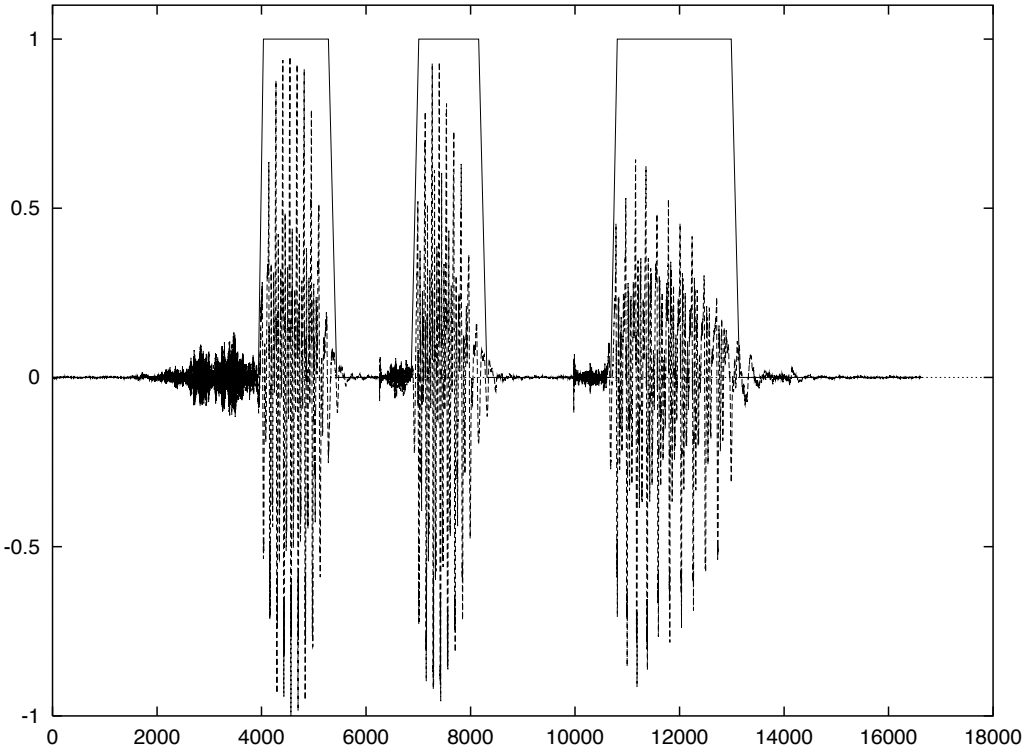

FIG. 8 – */SEUTEUTEU/ (locuteur "rg").*

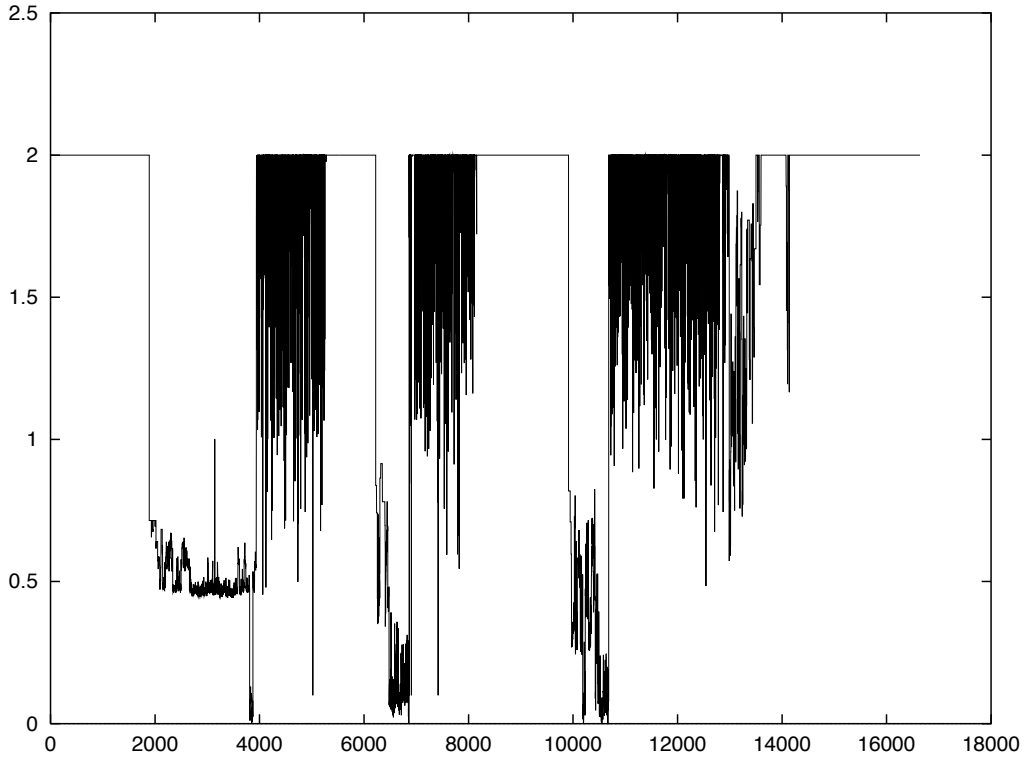

FIG. 9 – *Fonction de Hölder du locuteur "rg".*

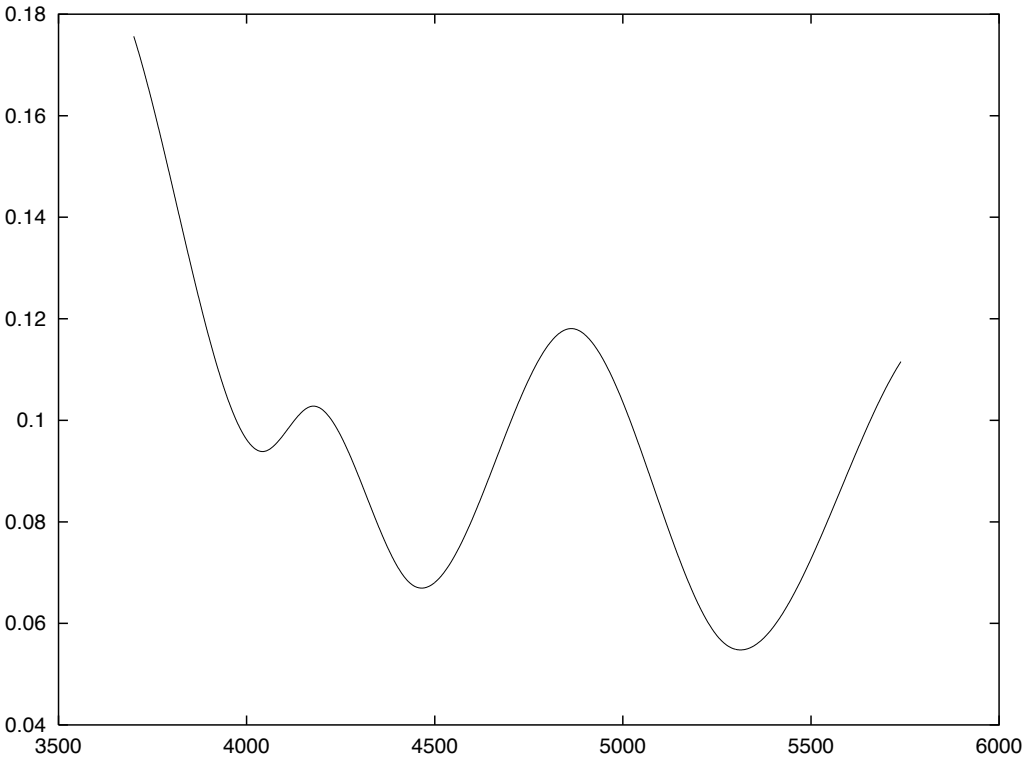

FIG. 10 – $\tilde{s}(t)$ *de la première partie non voisée(locuteur "ob").*

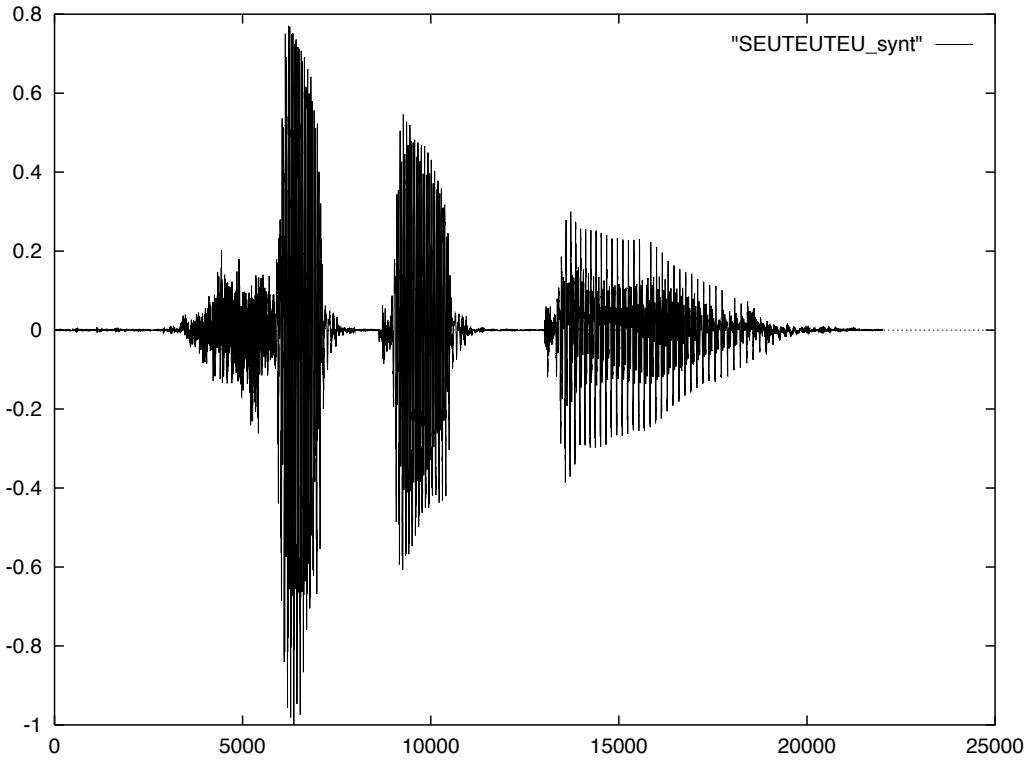

FIG. 11 – *SEUTEUTEU/ reconstruit (locuteur "ob").*

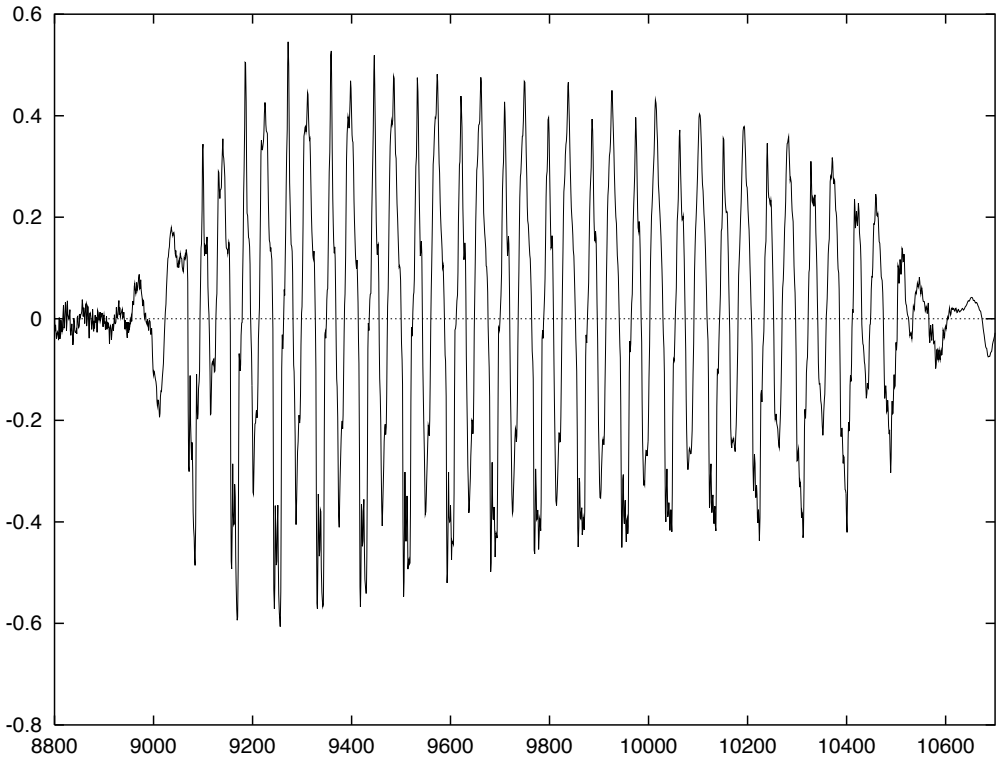

FIG. 12 – */TEU/ (locuteur "ob").*

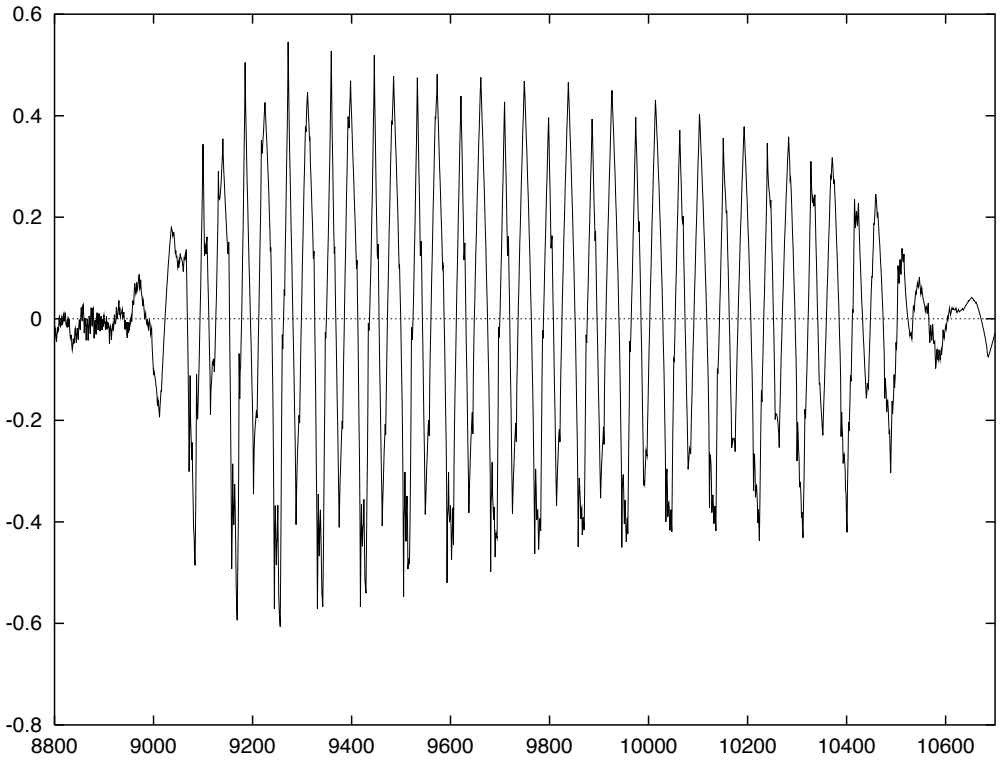

FIG. 13 – */TEU/ reconstruit (locuteur "ob").*

+ bruit d'un signal de parole. Le principe de ce modèle repose sur l'hypothèse que le signal est composé d'une partie déterministe et d'une partie stochastique. Une représentation harmonique est associée aux structures périodiques (parties voisées) du signal de parole et les parties "stochastiques" (i.e bruit de friction, turbulences glottales...) sont modélisées par un bruit. Plus précisément, les parties du spectre de Fourier allant jusqu'à la fréquence maximale de voisement sont modélisées par des sommes de composantes harmoniques avec des amplitudes complexes linéaires par morceaux. Quant à la partie "stochastique", elle est obtenue en filtrant un bruit blanc Gaussien par un filtre tout pôle. Ce modèle permet d'effectuer des synthèses de parole de qualité remarquable.

L'approche fractale consiste à conserver la modélisation de la partie harmonique, qui est très régulière, et à la soustraire au signal original. Nous obtenons ainsi un signal résiduel consitué exclusivement de hautes fréquences. Nous modélisons ensuite ce signal par la méthode décrite ci-dessus.

Les figures ci-dessous montrent les résultats obtenus pour $j_0 = 2$, c'est-à-dire, lorsque l'on garde un échantillon sur quatre. Sur cet exemple, les s_j obtenus à des échelles successives sont presque identiques. Ceci justifie *a posteriori* l'utilisation du modèle, et rend même la synthèse encore plus économique, puisqu'il suffit de garder les paramètres définissant $\tilde{s}_1$.

La figure 14 représente le signal résiduel original. Les figures suivantes représentent respectivement la fonction de Hölder du signal résiduel original, celle du signal résiduel original sous échantillonné d'un pas de 2, leur lissage par simple moyenne mobile et le signal résiduel reconstruit.

On peut aller jusqu'à $j_0 = 8$ (ce qui revient à garder un échantillon sur 8) tout en obtenant un signal de parole synthétique indiscernable de l'original. Les rapports signal sur bruit (RSB) pour différentes valeurs de j_0 sont :

- Pour $j_0 = 1$: RSB = 29.5 Db.
- Pour $j_0 = 2$: RSB = 27.0 Db.
- Pour $j_0 = 3$: RSB = 24.6 Db.

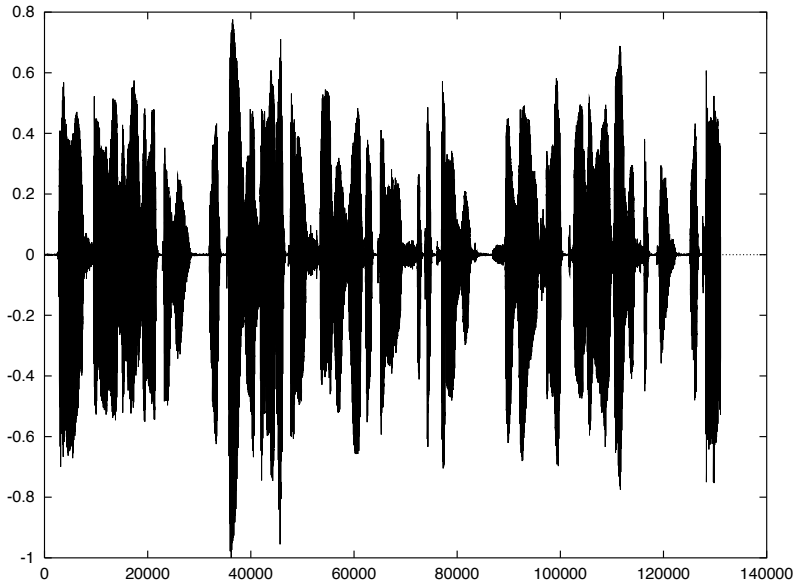

FIG. 14 – *Signal résiduel original.*

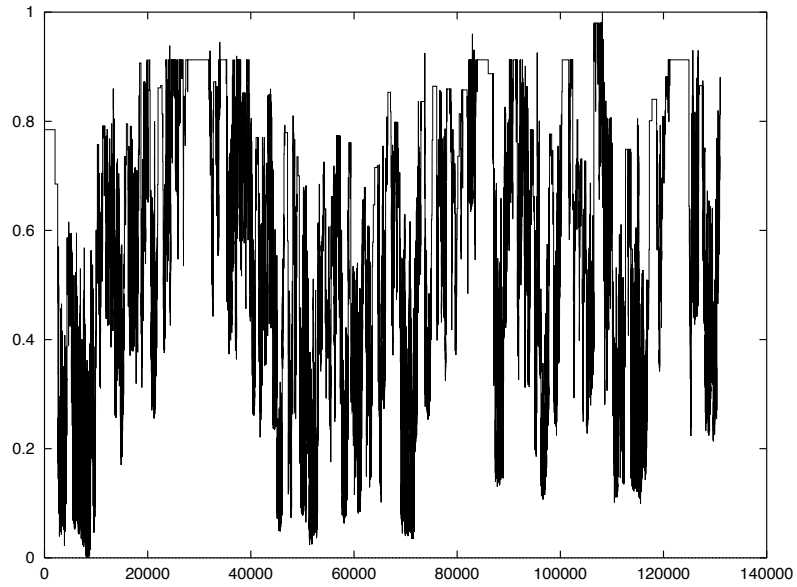

FIG. 15 – *Fonction de Hölder du signal.*

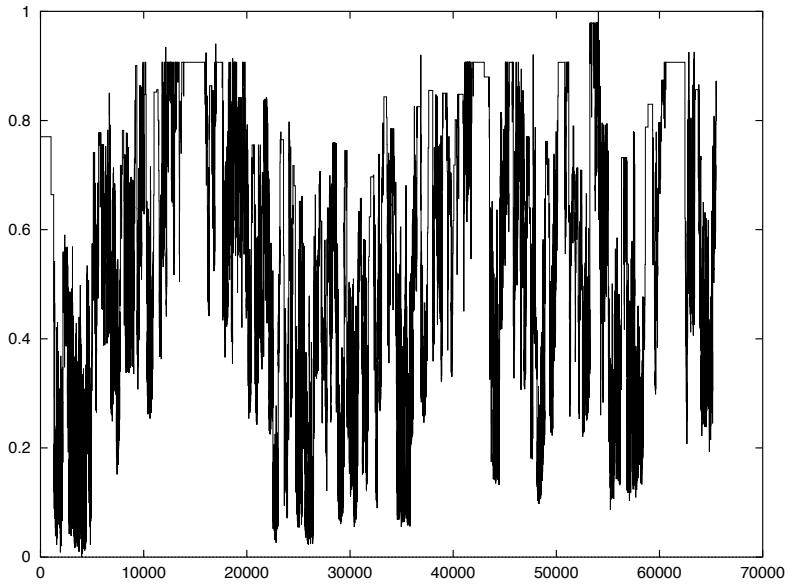

FIG. 16 – *Fonction de Hölder du signal signal à la résolution 2.*

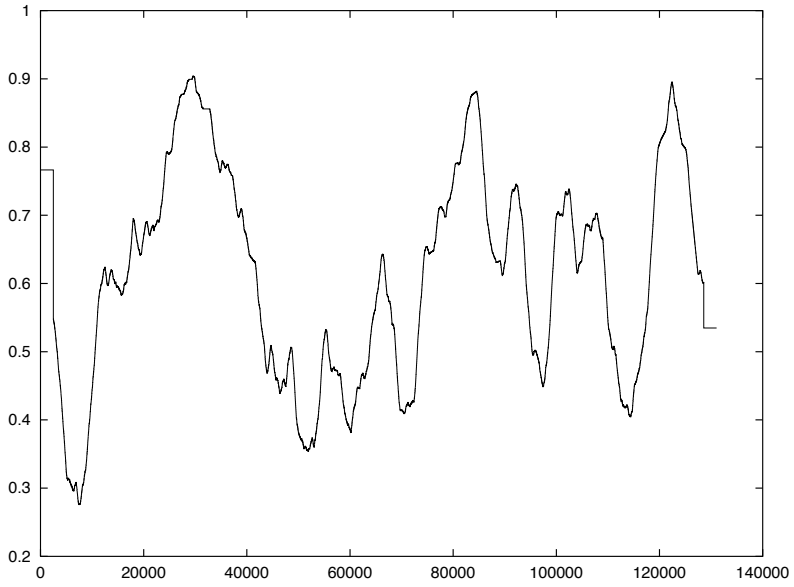

FIG. 17 – *Moyenne mobile sur la fonction de Hölder du signal* $(\tilde{s}_0)$.

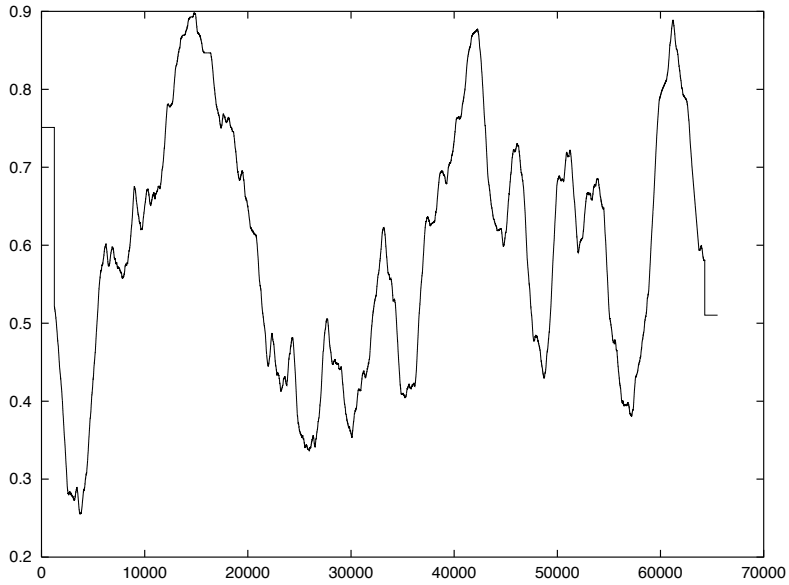

FIG. 18 – *Moyenne mobile sur la fonction de Hölder du signal à la résolution* $1/2$ $(\tilde{s}_1)$.

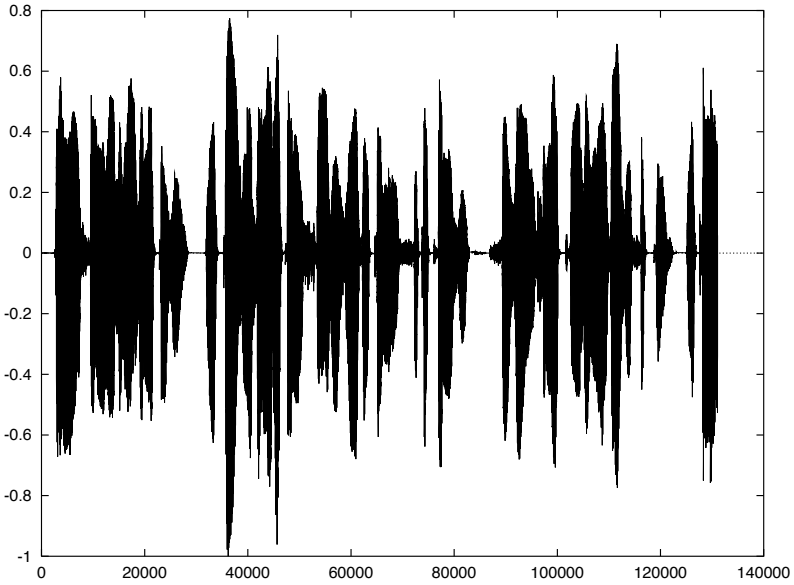

FIG. 19 – *Signal reconstruit.*

2 Analyse Multifractale et Traitement d'images

Dans de nombreux cas, l'information fournie par la fonction $\alpha(x)$ est soit difficile (voire impossible) à obtenir, soit trop complexe pour pouvoir être exploitée : par exemple $\alpha(x)$ est généralement partout discontinue pour les fonctions correspondant au graphe de l'attracteur d'un IFS. Il peut dans ce cas être avantageux de considérer une information simplifiée, de plus haut niveau, qui consiste à décrire, d'un point de vue géométrique ou statistique, la répartition des exposants de Hölder du signal : cette approche est appelée analyse multifractale, et trois type de spectres multifractals sont usuellement définis.

2.1 Spectre de Hausdorff

C'est ce spectre qui donne son nom à l'analyse multifractale, puisque l'on considère ici que chaque exposant de Hölder définit, comme ci-dessous, un ensemble fractal, dont on va calculer la dimension, et que le support du signal X est donc formé de l'union, éventuellement non dénombrable, d'ensembles fractals de dimension différentes. Plus précisément, soit :

$$E_\alpha = \{t/\alpha(t) = \alpha\}$$

On note

$$f_h(\alpha) = \dim_H E_\alpha$$

où $\dim_H$ dénote la dimension de Hausdorff. f_h est appelé spectre de Hausdorff de X ; il fournit une caractérisation géométrique des singularités du signal.

2.2 Spectre de grandes déviations

On essaie ici de caractériser de façon statistique les singularités.

Plus précisément, on définit, en se restreignant pour simplifier à un signal X défini sur $[0,1]$ et à une analyse par rapport aux intervalles dyadiques $I_n^k = [k2^{-n}, (k+1)2^{-n}[k = 0, \dots 2^n - 1, n \in \mathbb{N}$:

$$\alpha_n^k = -\frac{1}{n} \log |X((k+1)2^{-n}) - X(k2^{-n})|$$

α_n^k est appelé exposant de grain à la résolution n. Il mesure, en gros, la vitesse à laquelle le signal varie dans une période de temps de 2^{-n} : si $\alpha_n^k = 1$, la variation est "douce", plus α_n^k est petit, plus la variation est grande, et plus α_n^k est grand, plus la variation est petite.

On veut alors évaluer la probabilité de trouver un exposant donné quand on choisit un intervalle I_n^k au hasard dans $[0,1]$, ou, plus précisément, comment cette probabilité évolue quand n tend vers l'infini.

La double caractérisation (α, "probabilité" de trouver α) permettra alors de connaître tous les types de variations du signal ainsi que leur fréquence d'occurrence.

Pour cela, on pose :

$$f_g(\alpha) = \lim_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{\log N_n^\varepsilon(\alpha)}{n}$$

où

$$N_n^\varepsilon(\alpha) = \mathrm{card}\{k/(\alpha_n^k - \alpha) < \varepsilon\}$$

De façon informelle, cela revient à écrire que le nombre d'intervalles dyadiques de taille 2^{-n} ayant un exposant de grain à peu près égal à α varie comme $2^{n f_g(\alpha)}$, ou encore que :

$$P_n(\alpha_n^k \approx \alpha) \sim 2^{-n(1-f_g(\alpha))}$$

où P_n est la distribution uniforme sur tous les intervalles dyadiques de taille 2^n.

On voit immédiatement que $f_g(\alpha) < 1 \Rightarrow \lim_{n \to +\infty} P_n(\alpha_n^k \approx \alpha) = 0$. $f_g(\alpha)$ mesure alors la vitesse avec laquelle la probabilité de rencontrer α tend vers 0. f_g est appelé spectre de grande déviation, pour des raisons qui apparaîtront au paragraphe suivant. Notons que, de manière intuitive, les procédures qui conduisent au calcul de f_h et f_g sont très proches, à une "inversion de limite" près :

Pour f_h, on :

- calcule l'exposant de Hölder en faisant tendre la résolution vers l'infini,
- "compte" le "nombre" de points ayant un exposant fixé, c'est-à-dire qu'on calcule la dimension de Hausdorff de ces points.

Pour f_g, on :

- compte combien d'intervalles ont un exposant de grain fixé,
- fait ensuite tendre la résolution vers l'infini en effectuant une normalisation logarithmique pour obtenir f_g.

2.3 Spectre de Legendre

f_g s'interprète naturellement comme la fonction de taux d'un principe de grande déviation quand les hypothèses du théorème de Gätner–Ellis sont vérifiées. Voici les détails. On pose :

$$S_n(q) = \sum_{k=0}^{2^n-1} |X((k+1)2^{-n}) - X(k2^{-n})|^q$$

pour $q \in \mathbb{R}$, avec la convention $0^q = 0$.

On définit ensuite

$$\tau(q) = \liminf_{n\to\infty} \frac{\log S_n(q)}{-n}$$

Puis :

$$f_l(\alpha) = \tau^*(\alpha) = \inf_q(q\alpha - \tau(q))$$

f_l est appelé spectre de Legendre.

Posons alors : $Z_n = \log|X((k+1)2^{-n}) - X(k2^{-n})|$ où Z_n est une variable aléatoire et k suit la distribution uniforme P_n sur $\{0, \ldots, 2^n - 1\}$.

La fonction génératrice des moments correspondant à Z_n est :

$$C_n(q) = -\frac{1}{n}\log \mathbb{E}_n[e^{qZ_n}] = -\frac{1}{n}\log(2^{-n}S_n(q))$$

Le théorème de Gätner–Ellis stipule que, si $C(q) = \lim_{n\to\infty} C_n(q)$ existe (auquel cas cette limite vaut $1 + \tau(q)$) et est différentiable, alors :

$$C^*(\alpha) = f_g(\alpha) - 1$$

En d'autres termes :

$$f_l(\alpha) = f_g(\alpha)$$

Cette égalité est appelée formalisme multifractal faible.

2.4 Comparaison des spectres

Le spectre de Hausdorff fournit, on l'a vu, une caractérisation *géométrique* de la répartition des singularités du signal. f_g, lui, en décrit un aspect *statistique*. Enfin, f_l n'a pas de signification simple dans le cas général, mais, si le formalisme multifractal faible est valide, il fournit une méthode d'estimation de f_g beaucoup plus fiable que la définition directe : on remplace des calculs locaux et deux passages à la limite par un calcul global (une moyenne), et un seul passage à la limite.

Dans certains cas, en particulier si une structure multiplicative "stricte" est à l'œuvre, on dispose même d'un formalisme multifractal fort, qui indique que :

$$f_h = f_g = f_l$$

Ce cas est très favorable, car f_h est encore bien plus difficile à évaluer que f_g : calcul d'exposants ponctuels et de dimension de Hausdorff sont en effet en général très délicats aussi bien en théorie que dans les applications.

Dans le cas général cependant, seules les relations ci-dessous sont vraies sans aucune condition :

- $f_h \leq f_g \leq f_l$
- $f_l = f_g^{**}$
- $$\begin{array}{rcll} f_g(\alpha^+) & = & q\alpha^+ - \tau(q) & q > 0 \\ f_g(\alpha^-) & = & q\alpha^- - \tau(q) & q < 0 \end{array}$$

où $\alpha^+ = \tau'(q+), \alpha^- = \tau'(q-)$.

Ces relations montrent qu'en général, f_l contient moins d'information que f_g. En particulier, f_l est toujours concave.

2.5 Exemple : le mouvement Brownien fractionnaire (mBf)

Soit $X(t)$ un mBf d'exposant H, c'est-à-dire un processus gaussien de moyenne nulle et de fonction d'autocovariance :

$$\mathbb{R}(s,t) = E(X(s)X(t)) = \frac{\sigma^2}{2}[|t|^{2H} + |s|^{2H} - |t-s|^{2H}]$$

où $0 < H < 1$.

Les principales propriétés du mBf sont les suivantes :

- auto–affinité :

  $\forall a > 0\ X(at) \stackrel{d}{=} a^H\ X(t)$ où $\stackrel{d}{=}$ signifie égal en distribution.

- $Y_n^k = X((k+1)2^{-n}) - X(k2^{-n}) = X(2^{-n}) \stackrel{d}{=} N(0, 2^{-nH})$ i.e. les accroissements sont stationnaires et suivent une loi gaussienne centrée de variance 2^{-2nH}.

– "longue dépendance" :

$$S(\lambda) \sim |\lambda|^{1-2H} \text{ quand } \lambda \to 0$$

où $S(\lambda)$ est la densité spectrale de X.

– irrégularité locale :

$$\text{p.s., } \alpha(t) = H \ \forall t$$

En particulier, p.s.,

$$\begin{aligned} f_h(H) &= 1 \\ f_h(\alpha \neq H) &= -\infty \end{aligned}$$

Les propriétés spectrales du mBf sont donc déterminées par H à la fois aux basses et aux hautes fréquences.

– Spectres multifractals : ce sont bien sûr des fonctions aléatoires, et l'on a p.s.:

$$\begin{aligned} f_g(\alpha) = f_l(\alpha) = -\infty \quad &\text{pour} \quad \alpha < H \\ f_g(\alpha) = f_l(\alpha) = 1 \quad &\text{pour} \quad \alpha = H \\ f_g(\alpha) \leq f_l(\alpha) \leq 1 + H - \alpha \quad &\text{pour} \quad \alpha > H \end{aligned}$$

Définissons des versions déterministes de f_g et f_l :

$$\begin{aligned} T(q) &= \lim_{n\to\infty} \frac{\log E_\omega[S_n(q)]}{-n} \\ F_l(\alpha) &= \inf_q (q\alpha - T(q)) \\ F_g(\alpha) &= 1 + \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \frac{\log \pi_n^\varepsilon(\alpha)}{n} \\ \pi_n^\varepsilon(\alpha) &= P_\omega \times P_n(\alpha_n^k \in [\alpha - \varepsilon_1 \alpha + \varepsilon]) \end{aligned}$$

Alors :

$$\begin{aligned} F_g(\alpha) = F_l(\alpha) = -\infty \quad &\text{si} \quad \alpha < H \\ F_g(\alpha) = F_l(\alpha) = 1 + H - \alpha \quad &\text{si} \quad \alpha \geq H \end{aligned}$$

2.6 Analyse multifractale de mesures

Pour l'application aux images, ainsi qu'à divers autres domaines, il est plus agréable de travailler avec des *mesures* qu'avec des *fonctions*. En particulier, une mesure incorpore par définition la notion de résolution, qui doit par contre être spécifiée en sus si on travaille avec des fonctions. Nous exposons donc brièvement dans ce paragraphe l'analyse multifractale des mesures, qui fut d'ailleurs la première à être développée, et exposons en détail le cas de la mesure binomiale, archétype de mesure multifractale.

2.6.1 Exposant de Hölder

Soit donc μ une mesure de probabilité sur $[0,1]$. Comme plus haut, nous restreignons à des intervalles dyadiques, bien que des définitions plus générales puissent être posées.

L'exposant de grain de μ en $x \in (0,1)$ à la résolution n est :

$$\alpha_n(x) = \frac{\log \mu(I_n(x))}{-n}$$

où $I_n(x)$ est l'intervalle dyadique contenant x.

L'exposant de Hölder de μ en x est :

$$\alpha(x) = \lim_{\varepsilon \to 0} \frac{\log \mu(\beta(x,\varepsilon))}{\log \varepsilon}$$

quand la limite existe.

2.6.2 Spectre de Hausdorff

C'est exactement la même définition que plus haut :

$$\begin{aligned} E_\alpha &= \{x/\alpha(x) = \alpha\} \\ f_h(\alpha) &= \dim_H E_\alpha \end{aligned}$$

2.6.3 Spectre de grandes déviations

On pose encore

$$N_n^\varepsilon(\alpha) = \mathrm{card}\{k/|\alpha_n^k - \alpha| < \varepsilon\}$$

et

$$f_g(\alpha) = \lim_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{\log N_n^\varepsilon(\alpha)}{n}$$

2.6.4 Spectre de Legendre

$$\tau(q) = \liminf_{n\to\infty} \frac{\log \sum_{k=0}^{2^n-1} \mu(I_n^k)^q}{-n}$$

avec la convention $0^q = 0$

$$f_l(\alpha) = \tau^*(\alpha) = \inf_q(q\alpha - \tau(q))$$

Dans le cas des mesures, on a bien sûr aussi :

$$f_h \leq f_g \leq f_l$$
$$f_g^{**} = f_l$$

2.6.5 Exemple : la mesure binomiale

On définit la suite de mesure $(\mu_n)_{n\geq 1}$ par :

$$\mu_0 = \mathcal{L} \text{ (mesure de Lebesgue)}$$

$$\mu_n(I_n^k) = \mu_n[k2^{-n}, (k+1)2^{-n}[= m_0^{n\Phi_0^n} m_1^{n\Phi_1^n}$$

où $m_0 \in]0,1[, m_0 + m_1 = 1$ et Φ_0^n est la proportion de 0 dans le développement en base 2 de x quelconque dans I_n^k jusqu'à l'ordre n, et $\Phi_q^n = 1 - \Phi_0^n$

$$x = \sum_1^\infty x_i 2^{-i} \qquad x_i = 0, 1$$
$$\Phi_0^n(x) = \frac{\sum_{i=1}^n (1 - x_i)}{n}, \quad \Phi_1^n(x) = \frac{\sum_{i=1}^n x_i}{n}$$

La suite $(\mu_n)_{n\geq 1}$ a une limite faible notée μ, et appelée mesure binomiale.

Calculons $\alpha_n(x)$:

$$\alpha_n(x) = -\Phi_0^n(x) \log_2 m_0 - \Phi_1^n(x) \log_2 m_1$$

On montre que $\alpha(x)$ existe si et seulement si $\lim_{n\to\infty} \Phi_0^n(x) = \Phi_0(x)$ existe, et dans ce cas :

$$\alpha(x) = -\Phi_0(x) \log_2 m_0 - \Phi_1(x) \log_2 m_1$$

Les ensembles E_α sont les ensembles de points de $[0,1]$ ayant une proportion de 0 dans leur développement en base 2 définie et fixée. C'est un résultat classique que :

$$\dim_H\{x/\Phi_0(x) = \Phi_0\} = -\Phi_0 \log_2 \Phi_0 - \Phi_1 \log_2 \Phi_1$$

Le spectre de Hausdorff est donc donné par :

$$\begin{cases} \alpha(\Phi_0) &= -\Phi_0 \log_2 m_0 - \Phi_1 \log_2 m_1 \\ f_n(\alpha) &= -\Phi_0 \log_2 \Phi_0 - \Phi_1 \log_2 \Phi_1 \end{cases}$$

Notons en passant que l'ensemble des points de $]0,1[$ "sans" Φ_0 ($\Phi_0^n(x)$ ne converge pas) est de dimension de Hausdorff égal à 1 (mais bien sûr de mesure nulle). Remarquons d'autre part que, $\mathcal{L}$ - p.s., $\Phi_0(x) = \Phi_1(x) = \frac{1}{2}$. On vérifie que, pour $\alpha_m = -\frac{1}{2}\log_2 m_0 m_1$, $f_n(\alpha_m) = 1$.

Si nous considérons maintenant les exposants de grain, nous voyons que $\mathcal{L}$ - p.s., $\alpha_n^k \to \alpha_m$, et $f_g(\alpha)$ va mesurer à quelle vitesse $Pr(|\alpha_n^k - \alpha_m| > \varepsilon)$ tend vers 0 pour $\varepsilon > 0$ quand $n \to \infty$.

Pour n fixé, il y a exactement C_n^k intervalles I_n^j tels que $\Phi_0^n(x) = \frac{k}{n}$ pour $x \in I_n^j$. Ceci permet d'évaluer $N_n^\varepsilon(\alpha)$, qui vaut aussi C_n^k pour ε suffisamment petit et α proche de $\alpha(\Phi_0 = \frac{k}{n})$. En utilisant ensuite la formule de Stirling pour estimer C_n^k, on peut obtenir $f_g(\alpha)$. Cependant, le calcul exact est un peu pénible du fait de la double limite, et il est beaucoup plus rapide d'évaluer $f_l(\alpha)$ et d'utiliser les relations générales entre les 3 spectres :

Par définition,

$$\tau(q) = \liminf_{n\to\infty} \frac{\log_2 \sum_{k=0}^{2^n-1} \mu(I_n^j)^q}{-n}$$

or, $\mu(I_n^j) = m_0^k m_1^{n-k}$ exactement C_n^k fois, $k = 0 \ldots n$. Donc,

$$\begin{aligned} \textstyle\sum_{j=0}^{2^n-1} \mu(I_n^j)^q &= \textstyle\sum_{k=0}^{n} C_n^k m_0^{kq} m_1^{(n-k)q} \\ &= (m_0^q + m_1^q)^n \end{aligned}$$

et

$$\tau(q) = -\log_2(m_0^q + m_1^q)$$

Un calcul simple de transformée de Legendre montre alors que $f_l(\alpha) = \tau^*(\alpha)$ est égal à $f_h(\alpha)$. Comme on a toujours $f_h \leq f_l \leq f_g$, il s'ensuit que le formalisme multifractal fort est valide pour la mesure binomiale : $f_\alpha = f_g = f_l$.

Cependant, il suffit de considérer la somme ou la concaténation de 2 telles mesures pour obtenir des cas où : $f_h = f_g \overset{<}{\neq} f_l$ (voir figures 20 à 23).

2.7 Application : Traitement des images

L'analyse d'images consiste à enseigner à un ordinateur comment "voir" et interpréter le contenu de scènes digitalisées. Les applications sont nombreuses : robotique mobile, contrôle de qualité, imagerie médicale (par exemple détection automatique de pathologies), imagerie satellitaire (par exemple contrôle de l'évolution de la végétation), ... Dans la plupart des

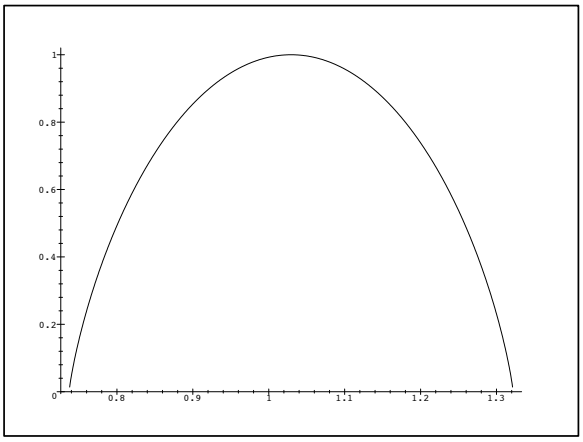

FIG. 20 – *Spectre d'une mesure binomiale.*

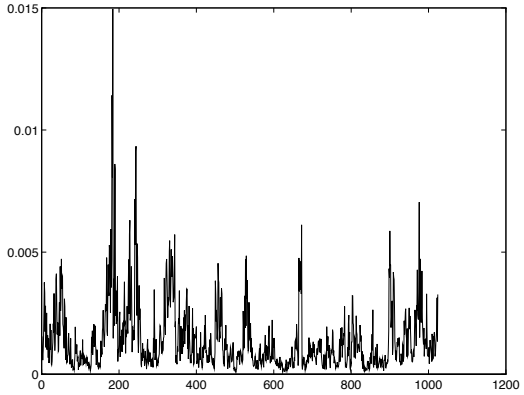

FIG. 21 – *Une mesure binomiale aléatoire.*

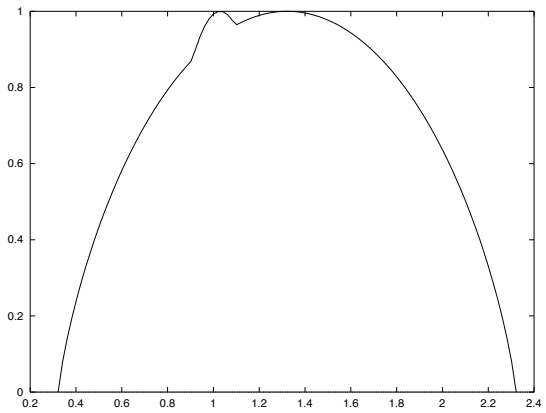

FIG. 22 – *Spectre de la concaténation de 2 mesures binomiales.*

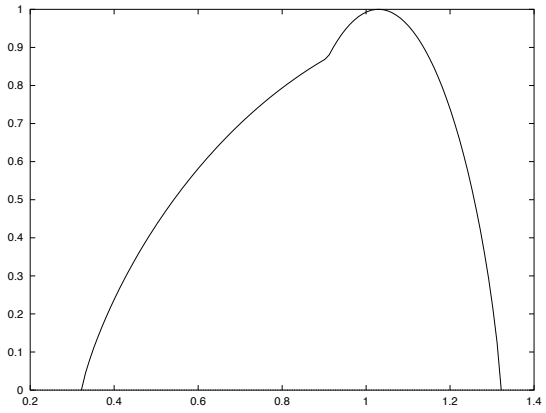

FIG. 23 – *Spectre de la somme de 2 mesures binomiales.*

cas, le but est de faire traiter par l'ordinateur des données qui sont en trop grand nombre pour pouvoir être exploitées manuellement (les images satellites en sont un bon exemple), ou bien d'être capable de détecter et de quantifier des caractéristiques que l'œil/le cerveau perçoit mal.

Une première étape dans la réalisation de ce programme est la détection et la localisation d'objets et de diverses autres caractéristiques dans des images. Un corpus considérable de recherches y a été consacré depuis vingt ans, permettant de développer des théories mathématiques élégantes et puissantes, et de résoudre certains des problèmes rencontrés. Plusieurs questions restent cependant ouvertes, aussi bien "à bas niveau" (segmentation d'images radar, par exemple) qu'à haut niveau (interprétation et reconstruction 3D par exemple). Parmi les approches désormais classiques en traitement d'images, citons la morphologie mathématique, les approches variationnelles, les modélisations probabilistes en particulier par champ de Markov, et l'analyse multi–échelles. Nous décrivons succintement ci-dessous les principes et la mise en œuvre de l'approche multifractale, en indiquant d'abord ses spécificités par rapport aux autres techniques.

2.7.1 Généralités

La plupart des approches "classiques" considèrent que l'image cible, c'est-à-dire le résultat à obtenir, est régulière, et que l'observation dont on dispose doit être "filtrée" (dans un sens vague) pour en extraire les éléments significatifs. Par exemple, les données peuvent être bruitées, et il s'agira de supprimer ce bruit par filtrage ou régularisation stochastique. Ou bien elles peuvent contenir des informations non pertinentes, comme par exemple des textures, qu'on lissera pour obtenir des régions uniformes.

A titre d'exemple extrêmement simple (on emploie aujourd'hui des techniques plus sophistiquée) qui permet de faire comprendre les idées sous-jacentes, considérons la détection de contours par filtrage de Canny. On se donne un modèle de contour, et on considère que l'image I observée est une version bruitée d'une image idéale contenant ce type de contours. Intuitivement, il paraît raisonnable, puisque les contours correspondent à de fortes variations des niveaux de gris, de les identifier aux points où la norme du gradient est maximale dans la direction du gradient. Comme les données sont :

1. discrètes,

2. bruitées,

on filtre I par une fonction C^∞, f, et on estime le gradient par $(I * f)' = I * f'$, avant de calculer la norme. En se donnant un modèle de bruit et des critères de qualité raisonnables, il est possible de dériver des filtres optimaux. Par exemple, pour une marche d'escalier avec un bruit blanc Gaussien, le filtre donnant la meilleure détection/localisation est :

$$f(x) = ae^{\alpha x} \sin ax + be^{\alpha x} \cos wx + ce^{-\alpha x} \sin wx + de^{-\alpha x} \cos wx$$

où α est un paramètre qui règle le degré de lissage. Cette approche peut être raffinée en donnant une version multi–échelles, qui devient particulièrement élégante et puissante dans le cadre des ondelettes. Nous n'allons pas passer en revue les nombreuses méthodes existantes, mais seulement noter les caractéristiques suivantes communes à la quasitotalité de ces méthodes :

- l'image originale est considérée comme contenant un "bruit" dont on veut se débarasser pour atteindre l'information pertinente,
- cette information est *de facto* supposée régulière, puisqu'elle est obtenue soit par filtrage, soit en effectuant une minimisation dans un espace de régularité,
- un paramètre exogène règle la force du lissage appliqué aux données originales.

L'approche multifractale adopte un point de vue différent : au lieu de chercher à obtenir une (ou des) versions lissées de l'image (éventuellement à différentes résolutions) de façon à se débarasser du "bruit", on essaie d'extraire de l'information des irrégularités de l'image, considérées comme pertinente. Les contours, par exemple, ne seront pas considérées comme des points où le gradient d'un hypothétique signal sous-jacent dérivable est maximal, mais comme des points où une irrégularité spécifique existe. De fait, même dans le cas d'une image très simple composée d'un carré noir sur un fond blanc, les contours sont des points de *discontinuité*, et le calcul de la régularité locale permettra de les détecter de façon aussi sûre, et peut-être conceptuellement plus directe, que l'estimation du gradient après lissage. Bien sûr, les approches par régularisation tirent leur intérêt d'une certaine robustesse au bruit, que l'on peut précisément évaluer. Mais il ne faut pas sous-estimer le rôle de la tradition dans leur emploi généralisé : les méthodes de l'analyse classique ont fait leurs preuves, elles sont bien connues, et il est naturel de les utiliser, en les adaptant, dans le cadre de la segmentation d'images. Cependant, avec le développement des outils fractals, une alternative apparaît, et nous allons voir que les résultats expérimentaux sont raisonnables, même si la théorie et les algorithmes doivent encore progresser.

L'approche fractale considère donc l'image à traiter comme aussi irrégulière que l'on veut, et tente d'extraire les caractéristiques recherchées par une étude fine des singularités, via l'analyse multifractale. Trois niveaux contribue à l'analyse de l'image : ponctuel (exposant de Hölder ponctuel), local (exposant local, exposants 2–microlocaux, corrélation multifractales) et global (spectre multifractal).

Voici quelques autres critères permettant de comparer approches fractales et "classiques" :

- comme dans les autres méthodes, un principe d'invariance par rotation/translation/changement d'échelle est respecté ;
- contrairement aux autres méthodes, l'analyse n'est pas locale : une information relative à toute l'image, ou toute une portion d'image, est utilisée pour analyser chaque point ;
- l'information de base n'est pas la valeur relative des niveaux de gris, mais la régularité Hölderienne, qui admet plusieurs variations (locale ou globale, par exemple). Cette hypothèse est justifiée dans le cas de scènes complexes ou la texture joue un rôle fondamental : sur la figure 24, où on a accolé deux fonctions de Weierstrass bi–dimensionnelles, une méthode classique détecterait plusieurs contours, alors que l'approche fractale fournira seulement une ligne verticale. Cet exemple extrême (des expériences sur des images naturelles sont montrées plus loin) indique que le choix d'une méthode dépend du type de résultats attendu et de la classe d'images considérées : il est d'ailleurs aujourd'hui généralement admis qu'une méthode de segmentation "universelle" n'existe pas.

2.7.2 Description multifractale des images

Commençons par deux points techniques :

- Nous faisons l'hypothèse que, pour toutes les images analysées :

$$f_h = f_g$$

  Bien sûr, une image réelle n'a pas de raison de vérifier le formalisme multifractal, mais il est possible de raffiner cette hypothèse en la localisant.

- Nous avons présenté plus haut l'analyse multifractale des *mesures*. En réalité, le bon cadre pour le traitement des images est celui des *capacités de Choquet*. Celles-ci sont

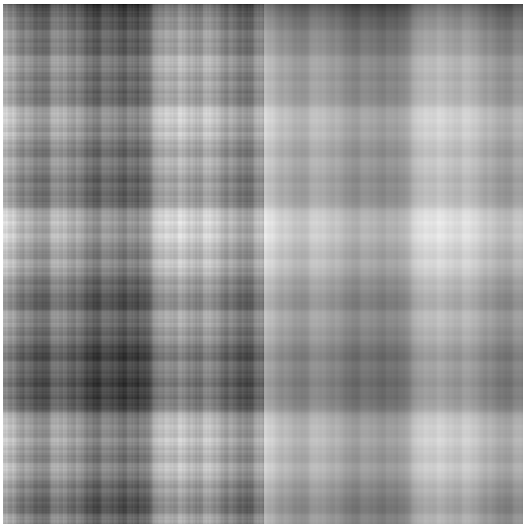

FIG. 24 – *Concaténation de 2 fonctions de Weierstrass bidimensionnelles.*

simplement des fonctions d'ensembles suffisamment régulières (comme les mesures), mais ne vérifiant pas nécessairement la propriété d'additivité, i.e.:

$$c(A \cup B) \neq c(A) + c(B) \text{ si } A \cap B = \emptyset$$

En effet cette restriction n'est d'aucune utilité pour l'analyse multifractale, et elle est artificielle dans le cadre du traitement d'images.

De plus, pour gagner encore un peu de souplesse, nous travaillerons avec des *suites* de capacités $(c_n)_{n\geq 1}$, n représentant la résolution d'analyse. Des exemples de capacités sont, outre les mesures usuelles, les suivants : on considère les intervalles dyadiques $I_n^k = [k2^{-n}, (k+1)2^{-n}[$ sur $[0,1]$, on pose $x_n^k = k2^{-n}$, et on se donne une suite de fonctions positives quelconques Θ_n de $P = \{I_n^k, k = 0 \ldots 2^n - 1, n \geq 0\}$ dans $[0,1]$.

Les fonctions suivantes définies sur les sous-ensembles de $[0,1]$ sont des suites de capacités de Choquet :

$$\begin{aligned} c_n(A) &= \max\{\Theta_n(I_n^k), x_n^k \in A\} \\ c_n^p(A) &= \left(\sum_{x_n^k \in A} \Theta_n(x_n^k)^p\right)^{\frac{1}{p}} \quad p \geq 1 \\ c_n(A) &= \max_t \text{ card}\{k/\Theta_n(I_n^k) = t, x_n^k \in A\} \end{aligned}$$

Remarque 2 *Si $A \subset I_n^k$, alors $c_n(A) = 0$ ou $c_n(A) = c_n(I_n^k)$.*
De telles capacités sont appelées "myopes".

Pour effectuer l'analyse multifractale de capacités, on pose :

$$\alpha_n(x) = \frac{\log c_n(I_n^k)}{\log \mu(I_n^k)}$$

où I_n^k est l'intervalle contenant x, et μ est une mesure de probabilité non atomique appelée mesure de référence. Les spectres sont définis de la même manière que pour les mesures.

Venons-en maintenant à l'analyse d'une image. Elle se déroule en 3 étapes :

- définition d'une suite de capacités à partir des niveaux de gris de l'image,
- calcul des exposants de Hölder et du spectre $f_g = f_h$,
- classification/traitement de chaque point x sur la base de l'information $(\alpha_x, f(\alpha_x))$.

α_x renseigne, de façon locale ou ponctuelle, sur la régularité de l'image au voisinage de x. Cependant, cette information n'est pas suffisante pour décider, par exemple, que l'on a affaire à un point de contour, et ce pour au moins 2 raisons :

- il n'existe pas de valeur intrinsèque de α que l'on puisse attribuer aux contours, ni même un ensemble de telles valeurs : les valeurs spécifiques obtenues pour une image donnée dépendent des conditions d'acquisition, ou d'une transformation non linéaire de l'image ;
- une information locale seule n'est pas suffisante pour décider si l'on a affaire à un point de contour. C'est le problème difficile des textures. Notons ici une forte divergence avec les méthodes de gradient, qui classent un point comme contour sur la base d'informations locales. La difficulté, et sa résolution, sont illustrées de façon schématique sur les figures 25 et 26 ci-dessous.

  La situation peut se résumer comme suit : s'il y a "trop" de "contours", l'œil aura plutôt tendance à parler de textures. Autrement dit, deux facteurs contribuent à faire classer un point comme contour :

  - la régularité locale : un contour correspond nécessairement à un certain type de singularité du signal, qui le distingue

– le "nombre" de points ayant ce même comportement dans une zone étendue autour du point étudié. Cette notion de "nombre" peut être rendue précise de deux manières, et c'est ici que nous utilisons l'hypothèse $f_h = f_g$: les contours lisses sont par définition des courbes de dimension $1 (f_h = 1)$. Ceci précise la quantité de points que l'on peut admettre comme point de contour d'un point de vue géométrique. Alternativement, l'interprétation statistique correspond au fait que lors du tirage au hasard d'un point dans l'image, un point de contour à une probabilité d'apparaître qui tend vers 0 quand la résolution tend vers l'infini avec une vitesse qui est précisément mesurée par f_g.

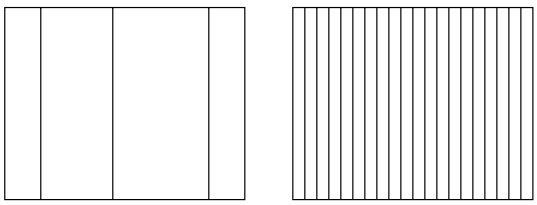

FIG. 25 – *Trois contours, une texture.*

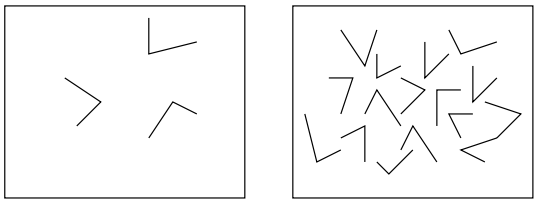

FIG. 26 – *Trois coins, une texture.*

En résumé, un point de contour est caractérisé par un α tel que :

– $f_h(\alpha) = 1$ car un contour lisse remplit l'espace comme une ligne ;

– $f_g(\alpha) = 1$ car un point de contour lisse a une probabilité donnée d'apparaître à une résolution fixée.

On peut pousser cette approche plus loin, et *définir* le "type" d'un point à travers la valeur associée de $f_h(\alpha)$:

Définition 2 *Soit $c = (c_n)$ une suite de capacités définies sur l'image. x est appelé point de c–contour si $f_n^c(\alpha(x)) = 1$. Plus généralement, x est appelé un point de c–type t, $t \in [0, 2]$ si $f_h^c(\alpha(x)) = t$.*

Remarquons que l'utilisation de l'information globale fourne par f_h requiert une certaine forme d'homogénéité, dans un sens multifractal. Cette contrainte peut cependant être affaiblie.

Venons en maintenant aux choix de la suite de capacités (c_n). Disons tout de suite qu'il n'exite pas de méthode générale pour choisir une suite qui soit optimale pour un problème donné. Les capacités utilisées habituellement tiennent compte de plusieurs critères :

- simplicité et rapidité de calcul,
- robustesse au bruit,
- pertinence par rapport au type d'image.

Trois capacités jouent un rôle particulier. On note Ω une région de l'image. $L(x, y)$ le niveau de gris en (x, y) :

- mesure somme :

$$c^s(\Omega) = \sum_{(x,y)\in\Omega} L(x, y)$$

- capacité max :

$$c^m(\Omega) = \max_{(x,y)\in\Omega} L(x, y)$$

- capacité iso :

$$c^i(\Omega) = \max_l \ \text{card}\{(x, y)/L(x, y) = l, (x, y) \in \Omega\}$$

On vérifie aisément que, alors que $c^s(\Omega)$ dépend à la fois des niveaux de gris et de leur répartition dans Ω, c^m dépend seulement des valeurs de niveaux de gris et c^i dépend seulement de leur répartition. Ce découplage des informations permet de raffiner l'analyse en utilisant des capacités vectorielles, les renseignements apportés par chaque coordonnée étant combinés via une approche Bayésienne.

Une fois les spectres associés à ces capacités calculés, on extrait par exemple les contours en détectant les points dont l'exposant α vérifie $f(\alpha) = 1$, comme expliqué plus haut. Un exemple de segmentation est donné figures 27 à 30.

Une autre application est la détection de changement dans des séquences d'images : ceci est par exemple important pour la surveillance de la terre par imagerie satellitaire, ou pour le contrôle de qualité. Une des difficultés est qu'il faut être capable de distinguer les changements fortuits (dus par exemple à un mauvais recalage des images ou à du bruit) des

FIG. 27 – *Image originale.*

FIG. 28 – *Egalisation d'histogramme sur l'image originale.*

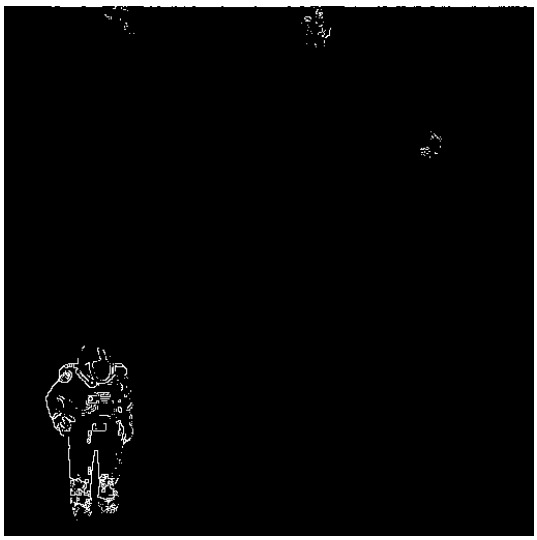

FIG. 29 – *Points de contours lisses obtenus avec la capacité max.*

FIG. 30 – *Points de contour irréguliers obtenus avec la capacité max.*

changements significatifs. Une possibilité dans notre cadre est de faire une analyse multifractale de l'image à l'instant t_{n+1} *par rapport* à une moyenne des images aux instants $t_n, t_{n-1} \dots t_1$.

Plus précisément, ayant défini une suite de capacités sur l'image à t_{n+1}, on en calcule le spectre multifractal en prenant comme mesure de référence μ la mesure somme d'une certaine moyenne des images précédentes. Si les images étaient strictement identiques, le spectre serait réduit à un point. En fonction du type de changement recherché (par exemple évolution naturelle de la végétation ou construction humaine), on choisira dans le spectre les zones d'intérêt en fonction de la dimension.

Une troisième application concerne le débruitage ou la restauration des images. Une méthode désormais classique consiste à seuiller les coefficients d'ondelettes s_j^k : si $|s_j^k|$ est inférieur à un seuil que l'on peut calculer en fonction de l'intensité du bruit dans l'image, on le remplace par 0, sinon on lui retranche ce seuil. Cette méthode donne en général de très bons résultats, et des théorèmes évaluent précisément sa validité dans diverses configurations, notamment en fonction du type de bruit.

Dans le cadre multifractal, on part de la remarque suivante : dans une image non bruitée, la plupart des points sont "lisses", c'est-à-dire ont un α égal à 2. Autrement dit, $f(2) = 2$. Dans une image bruitée, au contraire, la valeur presque sûre de α est strictement plus petite que 2. Une idée naturelle pour améliorer l'image est alors d'effectuer une transformation qui translate le spectre sur l'axe des abscisses de façon à retrouver $f(2) = 2$. Si on modifie de façon uniforme tous les exposants de Hölder, la force relative des singularités dans l'image sera conservée, et son aspect ne devrait pas être perturbé, à part la suppression du bruit. Un résultat de cette méthode sur une image radar est montré figures 31 à 33.

Pour effectuer en pratique une telle translation du spectre, on modifie les exposants de Hölder en chaque point en agissant sur la décroissance des coefficients d'ondelettes dans les échelles. La procédure exacte de décalage des coefficients est fondée sur une analyse 2–microlocale de l'image.

Références

[1] A. Arnéodo, E. Bacry, S. Jaffard, and J.F. Muzy. Oscillating singularities on cantor sets. To appear J. of Stat. Phys.

[2] D. Donoho. Denoising via soft thresholding. To appear IEEE Trans. Information Theory.

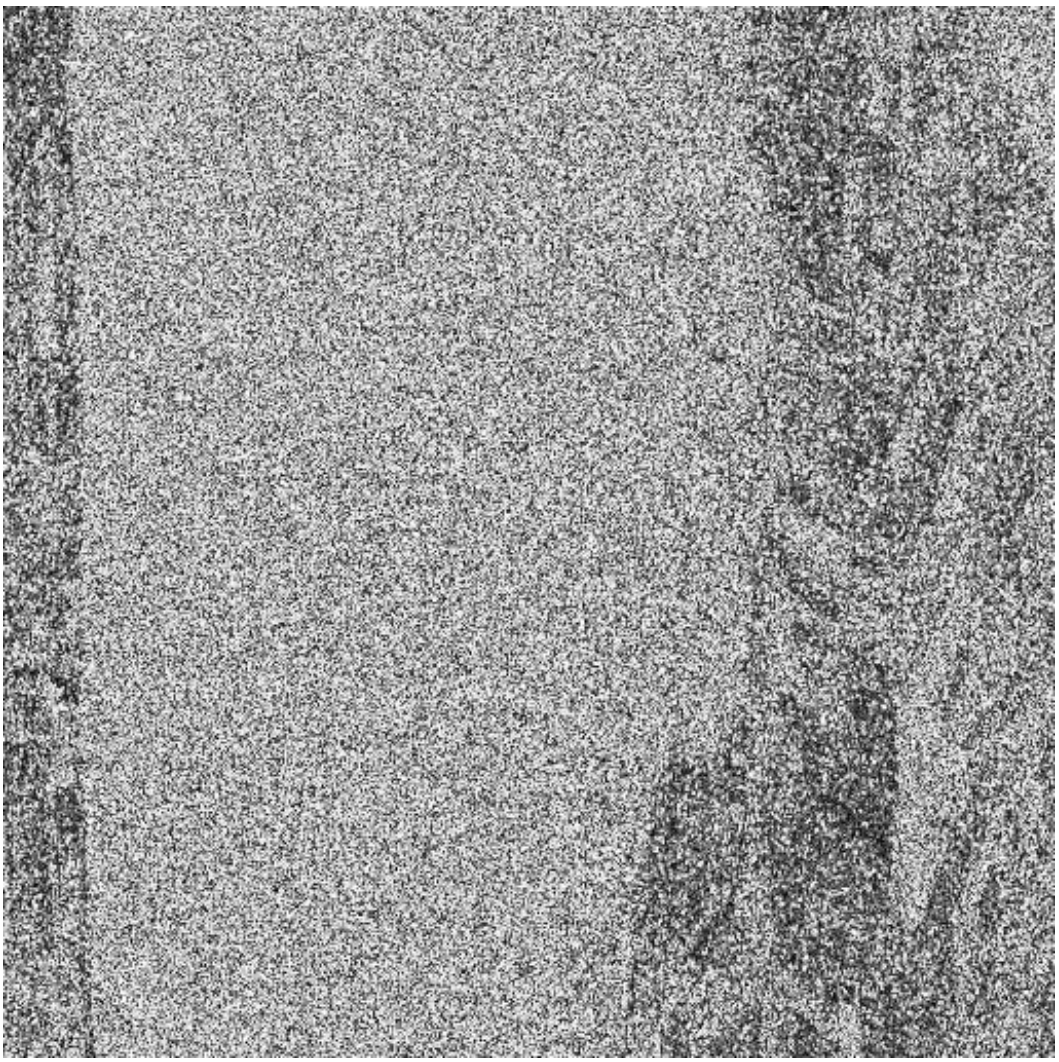

FIG. 31 – *Image radar originale.*

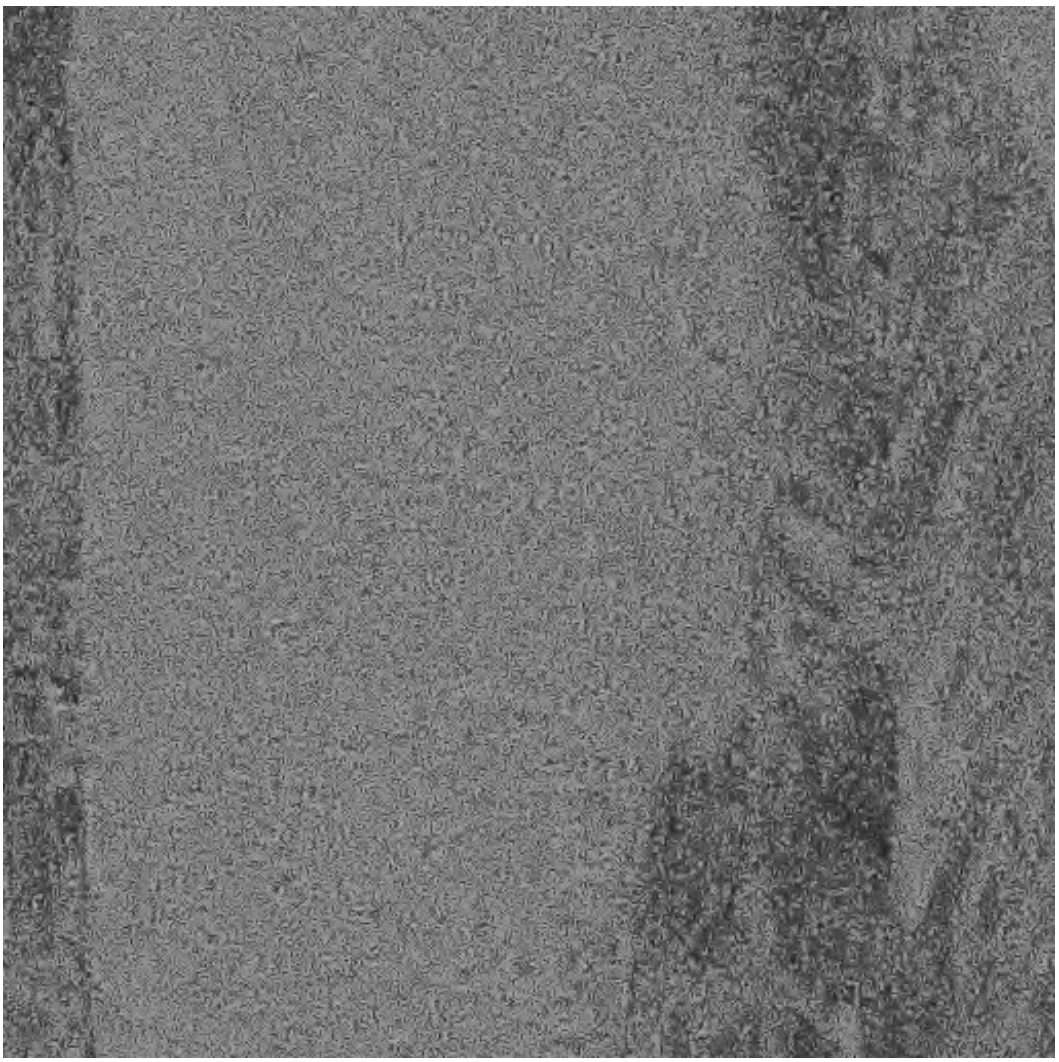

FIG. 32 – *Débruitage par seuillage des coefficients d'ondelettes.*

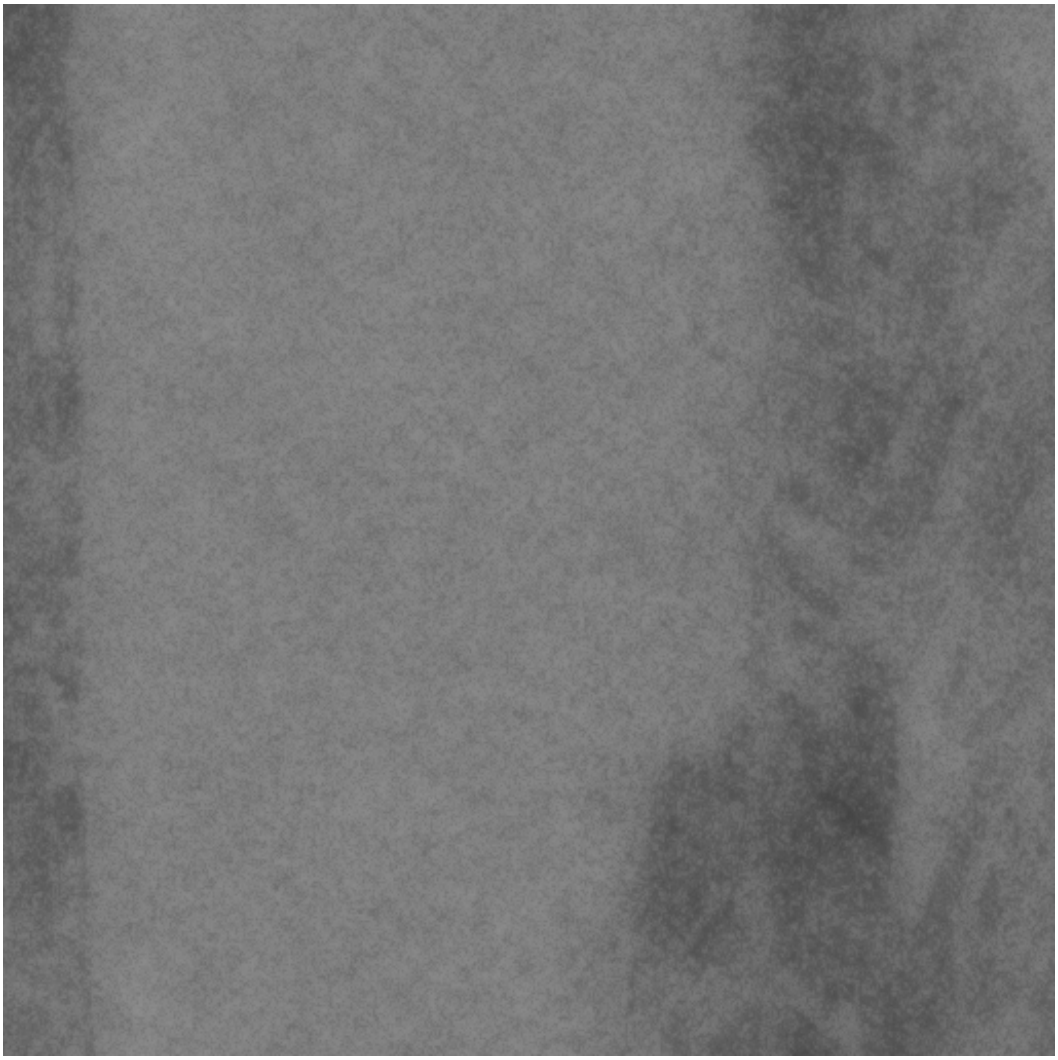

FIG. 33 – *Débruitage par décalage du spectre multifractal.*

[3] B. Guiheneuf and J. Lévy Véhel. Analyse 2–microlocale et traitement fractal des signaux. Technical report, INRIA, 1997.

[4] S. Jaffard and Y. Meyer. Wavelet methods for pointwise regularity and local oscillations of functions. *Memoirs of AMS*, 123(587), 1996.

[5] J. Lévy Véhel. *Fractal Image Encoding and Analysis*, chapter Introduction to the Multifractal Analysis of Images. Springer Verlag, 1997.