

Rapport de stage de M2 Mathématiques Avancées, parcours Probabilités et Statistiques

Algorithmes proximaux pour la segmentation de textures anisotropes

Leo Davy

supervisé par
Patrice Abry & Nelly Pustelnik

Soutenu le 2 Septembre 2022
devant le jury composé de
Alice Guionnet & Grégory Miermont



Laboratoire de Physique
École Normale Supérieure de Lyon
Mars-Juillet 2022



Table des matières

1	Introduction	2
1.1	Présentation des problèmes réels	2
1.2	Caractéristiques des textures	3
1.2.1	Choix des caractéristiques	3
1.2.2	Multi-échelle	4
1.2.3	Régularité	4
1.2.4	Anisotropie	5
1.2.5	Modèles	5
1.3	Analyse et segmentation	6
2	Modélisation de textures	9
2.1	Autosimilarité, régularité et images aléatoires	9
2.2	Stationnarité, covariance et variogramme	10
2.3	Champs autosimilaires anisotropes	11
3	Analyse de textures homogènes par ondelettes	16
3.1	Formalisme et analyse multi-échelle	16
3.2	Ondelettes pour l'analyse multi-échelle	17
3.3	Application sur AFBF et modèles	18
4	Optimisation et segmentation de textures hétérogènes	23
4.1	Estimation et segmentation	23
4.1.1	Stratégie générale	23
4.1.2	Attache(s) aux données et pénalisation	24
4.1.3	Modèles	26
4.2	Opérateurs proximaux	27
4.2.1	Théorie générale et formulation duale	27
4.2.2	Algorithmes proximaux	30
4.2.3	Problème dual	31
4.3	Forward-Backward et version accélérée	32
4.4	Algorithme primal dual et accélération par forte convexité	33
4.5	Résultats expérimentaux	35
4.5.1	AFBF hétérogène	35
4.5.2	Brodatz	42
5	Conclusion et perspectives	45

Chapitre 1

Introduction

Lors de ce stage, l'objectif a été d'introduire une méthodologie permettant la segmentation de textures caractérisées par leur irrégularité. On a en particulier cherché à pouvoir extraire des descripteurs directionnels dans les textures.

Un objectif a été d'avoir une vue d'ensemble du sujet afin de poursuivre l'étude du sujet en thèse à partir d'Octobre.

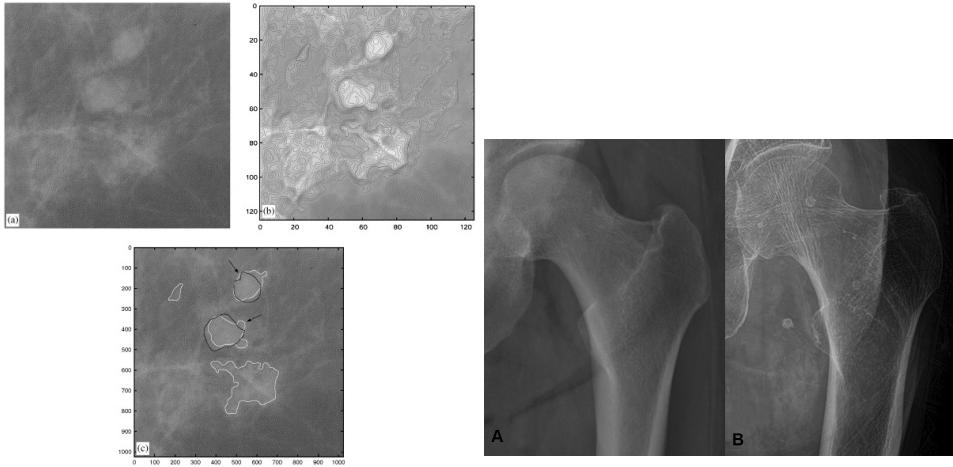
1.1 Présentation des problèmes réels

En traitement des images l'une des questions centrales est de pouvoir décrire ce qui compose une image. Dans ce contexte, les textures apparaissent comme des modèles naturels de descriptions élémentaires d'une grande partie des images, au moins localement. Dans de nombreuses applications réelles, il a été observé qu'une description des textures peut apporter des réponses aux questions posées (Cf. fig. 1.1). De plus, pour des problèmes plus élaborés tels que la vision assistée par ordinateur, la plupart des méthodes ont seulement été développées à partir de propriétés géométriques déterminées par des contours entre des textures lisses. Les textures auxquelles nous nous intéressons ont la particularité d'être caractérisée par des propriétés statistiques correspondant à une notion de régularité ou de rugosité.

Ainsi, le problème de segmentation de textures auquel nous nous intéressons ici peut à la fois avoir des applications directes sur des images réelles (par exemple en imagerie médicale) mais peut aussi être vu comme une capacité à enrichir la description des images de façon interprétable. Il faut donc dans un premier temps formaliser la notion de texture (modélisation, cf. Chapitre 2), puis développer une procédure d'extraction de caractéristiques (analyse, cf. Chapitre 3). L'analyse supposant une information arbitrairement précise il faut introduire des modèles permettant de compenser les approximations qui apparaissent toujours en pratique (optimisation, cf. Chapitre 4).

Avant de présenter la formalisation des modèles, considérons deux exemples sur lesquels, précisément, l'extraction d'attributs de textures est de grande importance :

1. (Fig. 1.1a) Pour les mammogrammes, les propriétés de régularité et d'anisotropie ont été identifiées comme indicateurs de lésions ou cancers [Ric15].
2. (Fig. 1.1b) La présence de stries (directionnelles) est un indicateur de l'ostéoporose.



(a) Mammogramme contenant des masses bénignes, puis contours d'iso-intensité obtenue par une analyse multi-échelle, puis version segmentée. Les flèches et les contours en noir indiquent les zones correctement détectées [RAL07].

(b) Radiographies de hanche, celle de droite est diagnostiquée ostéoporotique JANG et al. [Jan+21].

FIGURE 1.1

1.2 Caractéristiques des textures

1.2.1 Choix des caractéristiques

Dans ce travail, nous proposons de caractériser ce qu'est une texture à l'aide de trois concepts que nous allons rappeler puis relier entre eux.

1. Représentation multi-échelle : Depuis Mandelbrot[Man75], il est connu que la façon dont se comporte un objet quand on zoomé dessus donne de nombreuses propriétés. Mandelbrot a notamment pu mettre en évidence que le point de vue fractal/invariance d'échelle permettait de modéliser ou d'interpréter de nombreuses situations : par exemple, la distribution des galaxies, le relief terrestre ou bien le fromage suisse.
2. Régularité : Le concept de régularité ou de rugosité modélise la variation au voisinage d'un point. C'est une information naturelle à prendre en compte sur une image texturée. Cette approche nécessite cependant des modèles statistiques avancés pour décrire correctement cette notion. Par exemple, le concept de régularité est cohérent avec le modèle fractal aux échelles les plus fines.
3. Anisotropie : Dans des images réelles, contrairement à beaucoup d'études théoriques, c'est l'invariance par rotation qui relève de l'exception. Dans de nombreuses situations la directionnalité peut même être la propriété essentielle à observer. Ici, ce sera donc l'anisotropie qui relèvera du cas particulier et on cherchera à obtenir une information directionnelle la plus générale possible.

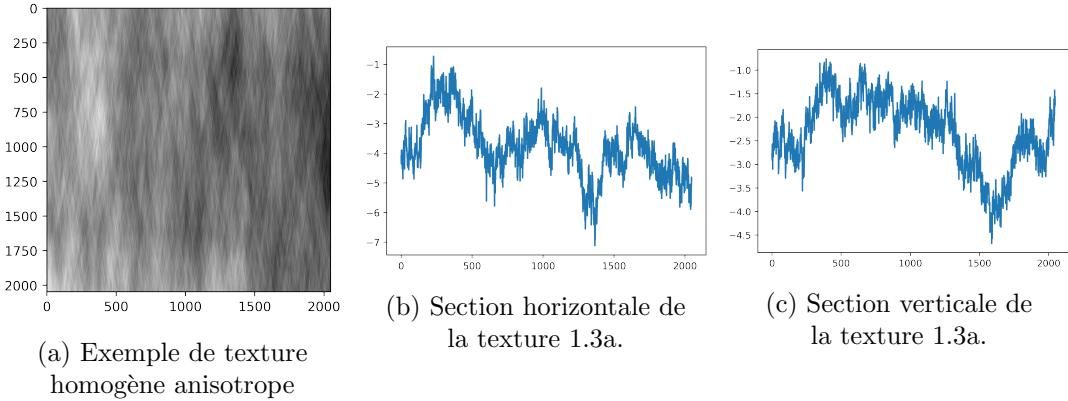


FIGURE 1.2 – Exemple de texture irrégulière anisotrope. Bien qu'à l'œil nu l'anisotropie soit évidente à observer, elle reste difficile à deviner à partir de sections horizontales et verticales de l'image.

1.2.2 Multi-échelle

Ces trois concepts introduits, nous allons maintenant commencer à les définir formellement et les relier entre eux.

Dans tout ce qui suit X désigne un signal, j une échelle¹ et $k = (x, y) \in \mathbb{R}^2$ une position. Généralement X est une image donc $X \in L^2(\mathbb{R}^2)$ dans le cas continu ou bien une version échantillonnée (discrétisée) du signal continu. Aussi, on s'intéressera à des quantités dites multi-échelles $T_X(2^j, k)$ qui seront définies par la suite à l'aide d'une transformée en ondelettes. De plus, quand $j \rightarrow -\infty$ on dit de façon interchangeable que l'on regarde aux échelles les plus fines ou aux plus hautes fréquences.

Les relations multi-échelles que nous considérerons seront celles d'une analyse fractale (Cf. partie 3.1)

$$T_X(2^j, k) \sim_{j \rightarrow -\infty} \sigma_k 2^{jh_k} \quad (1.1)$$

pour des signaux autosimilaires (Cf. partie 2.3)

$$\forall \lambda > 0 \quad X(\lambda \cdot) =^{f.d.d.} \lambda^h X(\cdot). \quad (1.2)$$

Il devient alors possible d'estimer la valeur de h et de $v = \log \sigma$ à la position k à partir d'une régression linéaire à travers les échelles j (Cf. fig. 1.3b). L'auto-similarité correspond au cadre formel dans lequel on peut réaliser l'analyse multi-échelle et c'est souvent la quantité d'intérêt par exemple dans [Man75]. En pratique cette quantité a notamment déjà été utilisée pour de l'analyse et de la classification de mammogrammes [Ric15] ou bien en physique expérimentale [Pas+18].

1.2.3 Régularité

La régularité à laquelle on s'intéresse est celle au sens de Hölder qui permet de généraliser les notions classiques de dérivées avec un point de vue fractionnaire. On dit que X est h -hölder

1. j correspond plus exactement à une octave, l'échelle correspondante est 2^j .

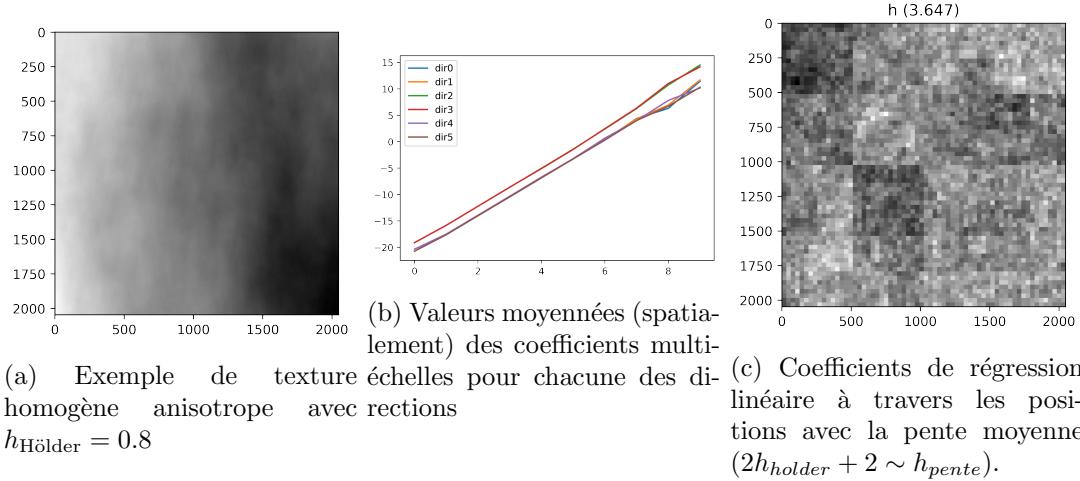


FIGURE 1.3

en k_0 s'il existe r suffisamment petit et un polynôme P de degré inférieur ou égal à h tel que pour tout $k \in B(0, r)$:

$$\|X(k_0 + k) - P(k)\| \leq C\|k\|^h. \quad (1.3)$$

L'analyse que l'on va réaliser suppose $h \in]0, 1[$ mais nous verrons avec l'introduction de la propriété des moments d'ondelettes nuls qu'il ne s'agit pas d'une condition restrictive.

1.2.4 Anisotropie

La prise en compte de l'anisotropie est la source principale de contributions de ce mémoire. On introduira (Cf. Chapitre 2) un point de vue stochastique spectral à l'aide d'une densité spectrale que l'on écrira

$$f(\xi) = \tau(\xi)\|\xi\|^{-(2h(\xi)+d)} \geq 0 \quad (1.4)$$

où τ et h sont deux fonctions π -périodiques définies sur \mathbb{S}_{d-1} . τ s'appelle *fonction de topothèse* et est à valeurs positives ; h est appelée *fonction de Hurst* et est à valeurs dans $(0, 1)$ (Cf. partie 2.3).

C'est grâce à ces deux fonctions que l'on induira, à partir d'une définition spectrale, des propriétés d'anisotropie et de régularité des textures. En effet, on verra qu'il est possible de générer des champs gaussiens à accroissements stationnaires avec une régularité de Hölder spécifiée à partir de

$$X^f(t) := \int_{\mathbb{R}^d} (e^{i\langle t, \xi \rangle} - 1) \sqrt{f(\xi)} \hat{W}(\mathrm{d}\xi) = \int_{\mathbb{R}^d} \sqrt{\tau(\xi)} \frac{(e^{i\langle t, \xi \rangle} - 1)}{\|\xi\|^{h(\xi)+d/2}} \hat{W}(\mathrm{d}\xi). \quad (1.5)$$

Ces champs sont appelés Champs Browniens Fractionnaires Anisotropes (AFBF - *Anisotropic Fractional Brownian Fields*).

1.2.5 Modèles

Sur ces champs, nous réaliserons une analyse à l'aide d'ondelettes complexes construites par analogie avec la transformée de Hilbert (Cf. partie 3.2). L'intérêt est que l'on dispose

alors d'une transformation avec de bonnes propriétés à la fois analytiques et numériques. Dans notre cas, l'une des propriétés majeures des ondelettes choisies est leur propriété de directionnalité d'un point de vue fréquentiel. On disposera alors de D ondelettes génératrices complexes $(\psi^{(d)})_{d=1}^D$ ce qui nous donnera D quantités multi-échelles (Cf. partie 3.3)

$$T_X^{(d)}(j, k) = |\langle X^f, \psi_{j,k}^{(d)} \rangle|^2 \sim |\langle f, \widehat{\psi}_{j,k}^{(d)} \rangle|^2, \quad (1.6)$$

où $\widehat{\psi}$ est la transformée de Fourier de ψ . Pour chacune de ces quantités on obtiendra alors une loi de puissance. En pratique, nous aurons $D = 6$ sur nos images qui proviennent d'une analyse horizontale, verticale et diagonale réalisée sur deux bandes duales.

Par régression log-linéaire à travers les échelles pour chacune des directions on extraira alors des paramètres de pente et d'ordonnée à l'origine.

Dans le cas autosimilaire, non nécessairement isotrope, nous aurons un modèle où la pente ne dépend pas de la direction. Ainsi, l'information directionnelle sera uniquement portée par les ordonnées à l'origine. Nous obtiendrons donc un modèle à $D + 1$ paramètres en chaque point (Cf. modèle *h-libre-min*)

$$\forall d = 1, \dots, D \quad \log T_X^{(d)}(j, k) = (2h + 2)j + v_d. \quad (1.7)$$

Nous verrons également que dans certains cas non-autosimilaires, l'anisotropie pourra être portée par les pentes. Nous obtiendrons alors un modèle à $2D$ paramètres en chaque point (Cf. modèle *h-lié-min*)

$$\forall d = 1, \dots, D \quad \log T_X^{(d)}(j, k) = (2h_d + 2)j + v_d. \quad (1.8)$$

1.3 Analyse et segmentation

La stratégie pour segmenter une image texturée comporte trois parties :

1. Calcul d'attributs analytiques locaux.
2. Régression linéaire des attributs.
3. Regroupement des attributs ayant les mêmes propriétés.

Traditionnellement, les trois étapes sont faites l'une après l'autre. Le problème principal posé par cette méthode est qu'isoler les étapes a tendance à supprimer de l'information et les erreurs vont avoir tendance à s'accumuler à travers les étapes (Cf. fig 1.4²).

La difficulté majeure à laquelle nous faisons face est l'hétérogénéité (Cf. partie 4.1.1). Là où la classification d'images homogènes de taille suffisamment grande ne pose quasiment aucun problème. La segmentation se faisant sur des mélanges de textures de tailles plus petites est une tâche plus complexe. L'objectif est donc de pouvoir détecter des textures différentes pour des supports aussi petits que possibles. On cherche ainsi à faire un compromis entre des modèles qui sont localement les plus pertinents possibles et homogènes sur les plus grandes surfaces possibles (Cf. partie 4.1.2).

Notre stratégie d'estimation repose sur une formulation variationnelle dans laquelle on va chercher à minimiser ce compromis entre une pénalisation statistique d'attache aux données

2. On pourra comparer aux résultats obtenus en fin de mémoire Cf. fig. 4.26

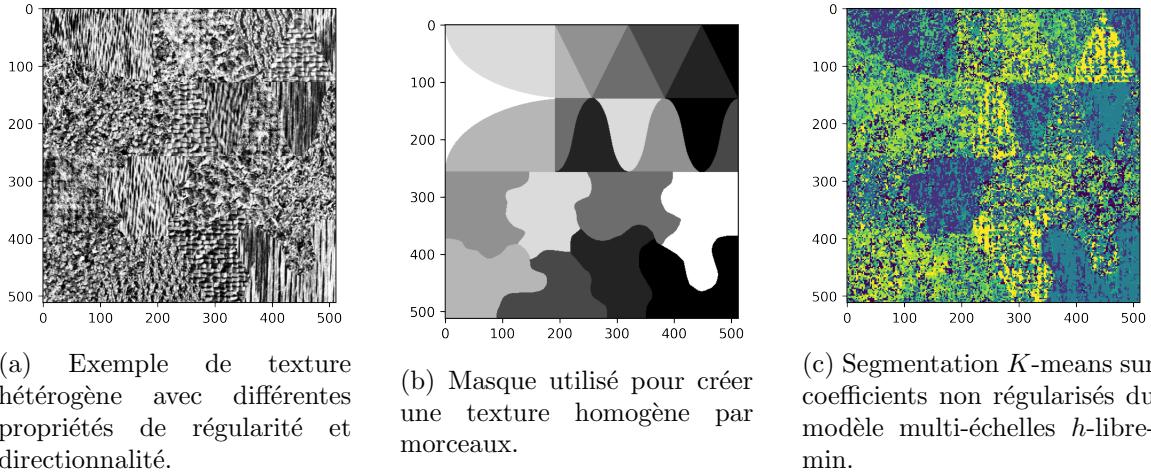


FIGURE 1.4 – Segmentation en trois étapes sur une mosaïque de Brodatz.

et une pénalisation géométrique d’homogénéité locale. Explicitement, le problème s’écritra (Cf. partie 4.1.3)

$$\text{minimiser}_{(V,H)} : \frac{1}{2} \|(V - V_{LR}, H - H_{LR})\|_{A^* A}^2 + \lambda T V_\alpha(V, H) \quad (1.9)$$

avec v un coefficient d’ordonnée à l’origine, h un coefficient de pente, (V_{LR}, H_{LR}) les coefficients de régression linéaire au sens des moindres carrés, A la matrice des variables explicatives et TV_α une généralisation de la variation totale.

Dans la fonctionnelle à minimiser, le premier terme correspond à l’attache aux données (à travers les échelles) tandis que la variation totale est un terme de régularisation des solutions (à travers l’espace).

On introduira alors la théorie des opérateurs proximaux (Cf. partie 4.2.1) qui permettent l’étude de fonctionnelles convexes non-lisses. Suite à cela, on pourra présenter des algorithmes permettant de résoudre les fonctionnelles d’intérêt (Cf. partie 4.2.2). On verra ensuite les expressions explicites d’algorithmes permettant la résolution d’un problème dual (Cf. partie 4.3) puis des algorithmes utilisant la structure des problèmes duals et primaux (Cf. partie 4.4). On verra à chaque fois des versions accélérées des algorithmes standard qui permettront d’obtenir de meilleures propriétés.

Finalement, on verra une section expérimentale dans laquelle on testera les méthodes de résolution proposées d’abord sur des réalisations d’AFBF (Cf. partie 4.5.1). On verra aussi une application sur des textures qui ne sont pas obtenues par des modèles mathématiques (Cf. partie 4.5.2).

Mentionnons également que pendant ce stage une bibliothèque Python a été développée qui a permis de vérifier empiriquement l’ensemble des méthodes proposées. Ainsi, toutes les images ont été obtenues à l’aide de notebooks Python. L’ensemble des codes peut être transmis sur demande, et sera mis à disposition en accès libre en ligne dès que cela sera possible.

On discutera alors en conclusion des contributions réalisées, notamment des modèles de prise en compte de l’anisotropie, ainsi que de leurs limites. On présentera alors différentes pistes de travail et de recherche qui seront explorées durant la thèse qui suivra ce stage.

Pendant cette thèse, on souhaitera d'une part approfondir les sujets qui n'ont pas pu l'être suffisamment pendant ce stage. D'autres part, on débutera l'étude de sujets dans le prolongement de ce qui a été étudié pendant ce stage et ayant le potentiel de repousser des limitations intrinsèques à l'approche présentée ici.

Chapitre 2

Modélisation de textures

2.1 Autosimilarité, régularité et images aléatoires

Dans ce chapitre nous allons nous concentrer sur l'introduction d'une grande classe de signaux (aléatoires) dont on pourra contrôler à la fois la régularité et des propriétés d'anisotropie.

Rappelons que l'objectif général de notre approche sera de pouvoir récupérer sur ces signaux des caractéristiques permettant de les distinguer. Dans le prochain chapitre nous verrons comment estimer ces caractéristiques de régularité et d'anisotropie localement. Dans ce chapitre, on peut travailler de façon globale et on parlera donc de texture homogène par la suite. Cependant, dans le contexte de ce chapitre il est plus naturel de parler d'accroissements stationnaires.

L'hypothèse de stationnarité nous permettra de compenser le fait que l'analyse de régularité traditionnelle se fait aux échelles les plus fines (auxquelles on n'a pas accès) par un moyennage spatial. Afin que cette méthode ait un sens il nous faudra donc une hypothèse sur les relations à travers les échelles, on parlera par la suite de formalisme fractal, et dans ce chapitre d'autosimilarité.

Les textures auxquelles nous nous intéressons sont issues de modèles stochastiques, c'est à dire que les textures sont obtenues par des processus aléatoires. Le modèle qui va être introduit a été choisi car il permet de réaliser des textures similaires à celles que l'on peut retrouver dans des champs applicatifs du traitement des images. De plus, les caractéristiques des réalisations sont relativement simples à définir dans un cadre analytique commode bien que abstrait.

Mentionnons en référence essentielle pour ce chapitre la thèse de Kévin Polisano [Pol17] et de Hermine Biermé [Bie05] dans lesquelles le cadre mathématique est défini précisément avec certains éléments mathématiquement plus approfondis que dans ce rapport. De plus, Kévin Polisano y explicite les méthodes de simulation des champs étudiés basé sur [BMR15]. Ce travail a notamment abouti à la bibliothèque Python PyAFBF[Ric22]¹ de Frédéric Richard qui nous permet de générer des réalisations des modèles présentés et qui nous permettra ensuite de tester les méthodes proposées.

1. <https://fjprichard.github.io/PyAFBF/index.html>

2.2 Stationnarité, covariance et variogramme

Le nombre d'échelles disponibles étant fini sur les applications réelles, une propriété souhaitée de nos modèles est celle de stationnarité.

Définition 2.2.1. On dit d'un processus $(X_t)_{t \in T}$ qu'il est :

- gaussien, si pour tout t_1, \dots, t_n on a que $(X_{t_1}, \dots, X_{t_n})$ est un vecteur gaussien.
- stationnaire, si pour tout $s, t \in T$, alors X_s et X_t ont la même distribution
- à accroissements stationnaires, si pour tout t_0 , on a que $(X_{t+t_0} - X_{t_0})_t$ a la même distribution que $(X_t - X_0)_t$

Dans ce qui suit, nous ne considérons que des champs gaussiens. Ainsi, nous avons les propriétés classiques de caractérisation par les deux premiers moments des distributions gaussiennes. De plus,

Proposition 2.2.2. Soit $(X_t)_{t \in T}$ un processus,

- Si $X = (X_t)_{t \in T}$ est stationnaire, alors, en particulier $\text{Cov}_X(s, t) = c_X(s - t)$ pour une certaine fonction c_X .
- (Prop 2.3 [Pol17]) Si $X = (X_t)_{t \in T}$ a des accroissements stationnaires, alors $\text{Cov}_X(s, t) = \frac{1}{2}(V_X(s) + V_X(t) - V_X(t - s))$ avec

$$V_X(t) = \mathbb{E}((X_t - X_0)^2) \quad (2.1)$$

appelé variogramme associé à X . De plus, il est équivalent de caractériser un champ gaussien à accroissements stationnaires par sa fonction de covariance ou son variogramme.

On peut observer dans la proposition, que la définition du variogramme $V_X(k)$ est particulièrement naturelle de notre point de vue d'étude de la régularité car par définition c'est la variation quadratique moyenne associée à X pour un accroissement k . On proposera notamment des illustrations de variogrammes (Cf. par exemple fig. 2.3 ou fig. 4.5) avec certaines réalisations afin de constater facilement les propriétés statistiques directionnelles des champs.

Voyons un premier exemple de champ gaussien, le mouvement brownien :

Définition 2.2.3. [LG13] On appelle mouvement brownien, tout processus mesurable $(B_t)_{t \in \mathbb{R}_+}$ à valeurs dans \mathbb{R} vérifiant :

1. $B_0 = 0$ p.s.
2. Le processus est à accroissements indépendants (pour tout $t_1 \leq t_2 \leq t_3$, on a que $B_{t_2} - B_{t_1}$ est indépendant de $B_{t_3} - B_{t_2}$).
3. $\forall s \leq t$ les accroissements suivent une loi normale $B_t - B_s \sim \mathcal{N}(0, |(t - s)|)$
4. Le processus est à trajectoire continues

Proposition 2.2.4. Soit $X = (X_t)_t$ un mouvement brownien, alors, son variogramme est donné par

$$V_X(t) = |t|. \quad (2.2)$$

Introduisons maintenant les définitions des propriétés que nous chercherons à identifier par la suite :

Définition 2.2.5. Soit X un champ gaussien sur \mathbb{R}^n , alors on dit que X est

1. auto-similaire d'ordre α , si pour tout $\lambda \in \mathbb{R}^+$ on a $X(\lambda \cdot) =^{f.d.d.} \lambda^\alpha X(\cdot)$.
2. isotrope, si pour toute rotation R de \mathbb{R}^n on a $X \circ R =^{f.d.d.} X$.
3. anisotrope, si X n'est pas isotrope.
4. h -Hölder régulier si il existe $C > 0$ tel que $\forall s, t$ on $|X(t) - X(s)| \leq C||t - s||^h$.

On obtient alors les propriétés suivantes pour le mouvement brownien :

Proposition 2.2.6. Proposition 2.5 et 2.11 [LG13]

1. Soit (B_t) un mouvement brownien standard, alors, pour tout $c > 0$, $\tilde{B}_t = c^{\frac{1}{2}} B_{\frac{t}{c}}$ est un mouvement brownien standard. Autrement dit, le mouvement brownien est auto-similaire d'ordre 1/2.
2. Le mouvement brownien est isotrope.²
3. Le mouvement brownien est 1/2-Hölderien.

2.3 Champs autosimilaires anisotropes

On va maintenant introduire le théorème de Bonami-Estrade qui permet de caractériser les symétries d'un champ gaussien à partir des propriétés d'une fonction f , appelée densité spectrale. La justification de l'utilisation d'un modèle spectral provient du théorème de Bochner (Théorème 2 [Bie19]) tandis que les résultats de régularité de Hölder proviennent du lemme de Kolmogorov-Chentsov (Théorème 4 [Bie19]). Le lecteur intéressé pourra en trouver les détails dans la thèse de Kévin Polisano [Pol17]³.

Définition 2.3.1. Soit $t \mapsto g(t)$ une fonction, on dit que g est

1. α -homogène si $\forall \xi, \forall \lambda > 0, g(\lambda \xi) = \lambda^\alpha g(\xi)$.
2. radiale si $\forall t, g(t) = g(||t||)$

Théorème 2.3.2. (Bonami-Estrade[BE03]) Soit f vérifiant

$$\int_{\mathbb{R}^2} \min(1, ||\xi||^2) f(\xi) d\xi < +\infty \quad (2.3)$$

et X^f un champ donné par sa représentation harmonisable

$$X^f(k) := \int_{\mathbb{R}^2} \left(e^{i\langle k, \xi \rangle} - 1 \right) f^{\frac{1}{2}}(\xi) \hat{W}(\xi). \quad (2.4)$$

où \hat{W} désigne la transformée de Fourier d'un bruit blanc gaussien.

Alors le champ est continu, et de plus,

— X^f est autosimilaire si et seulement si f est homogène.

2. En dimension 1 le groupe des rotations est seulement $\{-1, 1\}$, d'où le résultat par symétrie des distributions gaussiennes centrées.

3. En particulier, la définition de \hat{W} la transformée de Fourier d'une mesure de Wiener page 61 equation 1.22 et page 91 pour le théorème à proprement parler.

— X^f est isotrope si et seulement si f est radiale.

Finalement, si f vérifie

$$\exists A, B > 0 : f(\xi) \leq B||\xi||^{-(2h_0+2)} \quad \text{pour presque tout } |\xi| \geq A, \quad (2.5)$$

alors X^f est h_0 -Hölder régulier.

Introduisons maintenant la définition suivante afin de préciser les champs et notations utilisés pour tout ce qui suit.

Définition 2.3.3. Les densités spectrales f considérées dans ce rapport s'écrivent sous la forme

$$f(\xi) = \tau(\arg(\xi))||\xi||^{-(2h(\arg(\xi))+2)}. \quad (2.6)$$

où τ est appelée fonction de topothèse et h est appelée fonction de Hurst. Ces deux fonctions sont supposées définies sur \mathbb{S}_1 et π -périodiques, de plus τ est positive et H à valeurs dans $]0, 1[$.

L'intérêt de cette définition est de permettre de toujours vérifier la condition du Théorème 2.3.2 et de plus on sépare les conditions de radialité et d'homogénéité sur h et τ . En effet, changer τ n'influe pas la propriété d'homogénéité mais permet de contrôler la présence ou l'absence de radialité. Tandis que la présence ou l'absence d'homogénéité est contrôlée par le comportement de h .

Définition 2.3.4. Pour une densité spectrale f comme dans la Définition 2.3.3, on appelle X^f défini par 2.4 un champ brownien fractionnaire anisotrope (ou AFBF - Anisotropic Fractional Brownian Field).

Remarque 2.3.5. Implicitement, la notation X^f implique toujours que f est défini comme dans la Définition 2.3.3 et que X^f est défini comme dans le Théorème 2.3.2.

Finalement, on peut résumer les différents comportements de X^f sous la forme du corollaire suivant :

Corollaire 2.3.6. Soit X^f défini par la Définition 2.3.4, alors

1. X^f est autosimilaire si et seulement si h est constante.
 2. X^f est isotrope si et seulement si τ et h sont constantes.
 3. X^f est autosimilaire anisotrope si et seulement si τ n'est pas constante et H est constante.
- De plus, l'exposant de Hölder critique h_0 de X^f est donné par $h_0 = \text{ess inf}_{\xi} h(\xi) > 0$.

En particulier, à constante multiplicative près, les seuls champs auto-similaires isotropes ont des densités spectrales de la forme

$$f(\xi) = ||\xi||^{-(2h+d)}. \quad (2.7)$$

On peut alors choisir une extension de la définition du mouvement brownien avec un paramètre d'autosimilarité spécifié en dimension quelconque :

Proposition 2.3.7. Le champ brownien fractionnaire, en dimension d , est l'unique champ gaussien à accroissements stationnaires isotrope autosimilaire avec un paramètre d'autosimilarité $H \in]0, 1[$, sa densité spectrale est donc donnée par $f(\xi) = ||\xi||^{-d-2H}$.

En particulier, pour $H = \frac{1}{2}$, on parle de champ de Levy-Chentsov. Contrairement au mouvement brownien ($d = 1$, Cf. fig 2.1⁴) les accroissements ne sont plus indépendants en général. On a donc une classe de champs sur lesquels on peut choisir la régularité des réalisations et qui est paramétré par un unique nombre H . Remarquons que c'est cette classe de champs qui a été étudiée par Barbara Pascal dans sa thèse [Pas20] dans le même objectif de segmentation qu'ici. Le travail réalisé ici peut donc être vu comme une contribution anisotrope au travail de Barbara Pascal.

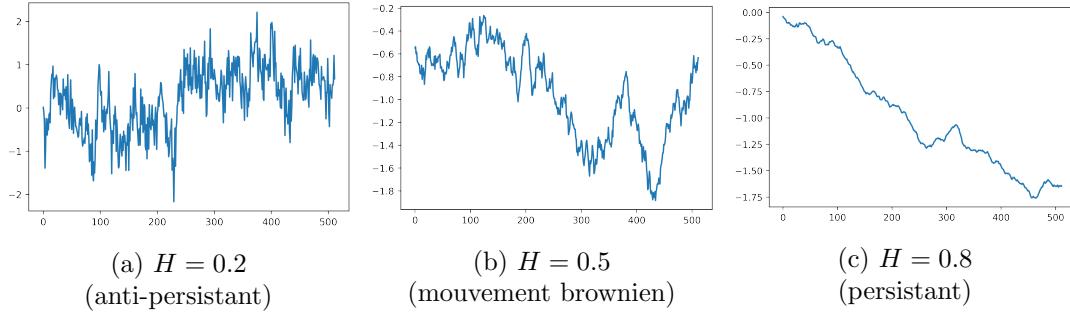


FIGURE 2.1 – Exemples de mouvements browniens fractionnaires.

Les premiers champs anisotropes autosimilaires que l'on introduit sont les *champs élémentaires* qui sont définis par une fonction de Hurst constante et une fonction τ de topothèse qui est une indicatrice. Du point de vue de l'interprétation, le support de τ agit comme un filtre fréquentiel qui supprime toutes les fréquences hors de son support. Les textures obtenues ont une directionnalité très nette (d'autant plus si le support est petit).

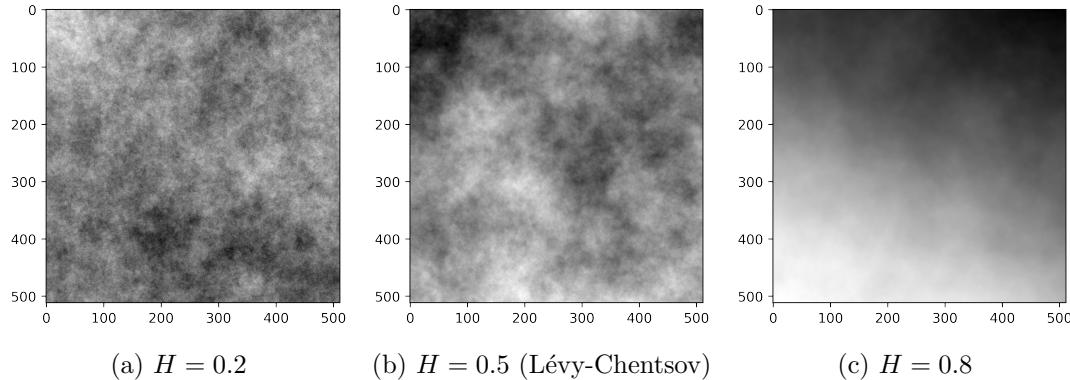


FIGURE 2.2 – Exemples de champs browniens fractionnaires.

Les champs fractionnaires anisotropes que l'on utilisera en pratique sont définis par des fonctions de Hurst h et de topothèse τ constantes par morceaux prenant deux valeurs, pour Hurst dans $]0, 1[$ et pour τ dans $]0, +\infty[$ (Cf. par exemple fig. 2.3).

Du point de vue de l'interprétation, la fonction de Hurst va spécifier la régularité des réalisations (proche de 0 : très irrégulier ; proche de 1 : assez lisse) tandis que la fonction τ va

4. Les processus de la figure 2.1 sont obtenus comme section des images de la figure 2.2. Il y a donc un léger abus de langage car les processus ne partent pas de 0.

pouvoir agir comme un filtre. Contrairement aux champs élémentaires, ici, τ peut augmenter ou diminuer l'importance des basses fréquences. Par contre, aux échelles les plus fines, l'influence de τ est négligeable sauf dans les directions où la fonction de Hurst est minimale [Ric18].

Cependant, les fonctions τ et de Hurst n'étant en général pas constantes, les champs ne sont en général ni autosimilaires, ni isotropes. Ces champs ont été choisis car ils offrent une paramétrisation assez simple et ont suffisamment d'expressivité pour ressembler à de nombreuses textures observées en pratique.

Afin de tester (Cf. section 4.5.1) les méthodes de segmentation introduites nous aurons besoin de textures hétérogènes. Le choix a été fait dans un premier temps d'utiliser une approche très simple en mélangeant deux textures homogènes normalisées à l'aide d'un masque. Cette méthode a comme inconvénient majeur le fait qu'une zone de discontinuité est introduite au niveau de l'interface entre les deux textures. Cela facilite en fait la segmentation car les zones de changement de textures sont alors assez faciles à identifier. Cependant on verra que les méthodes introduites permettent bien de distinguer les propriétés statistiques des textures sur chacune des zones homogènes.

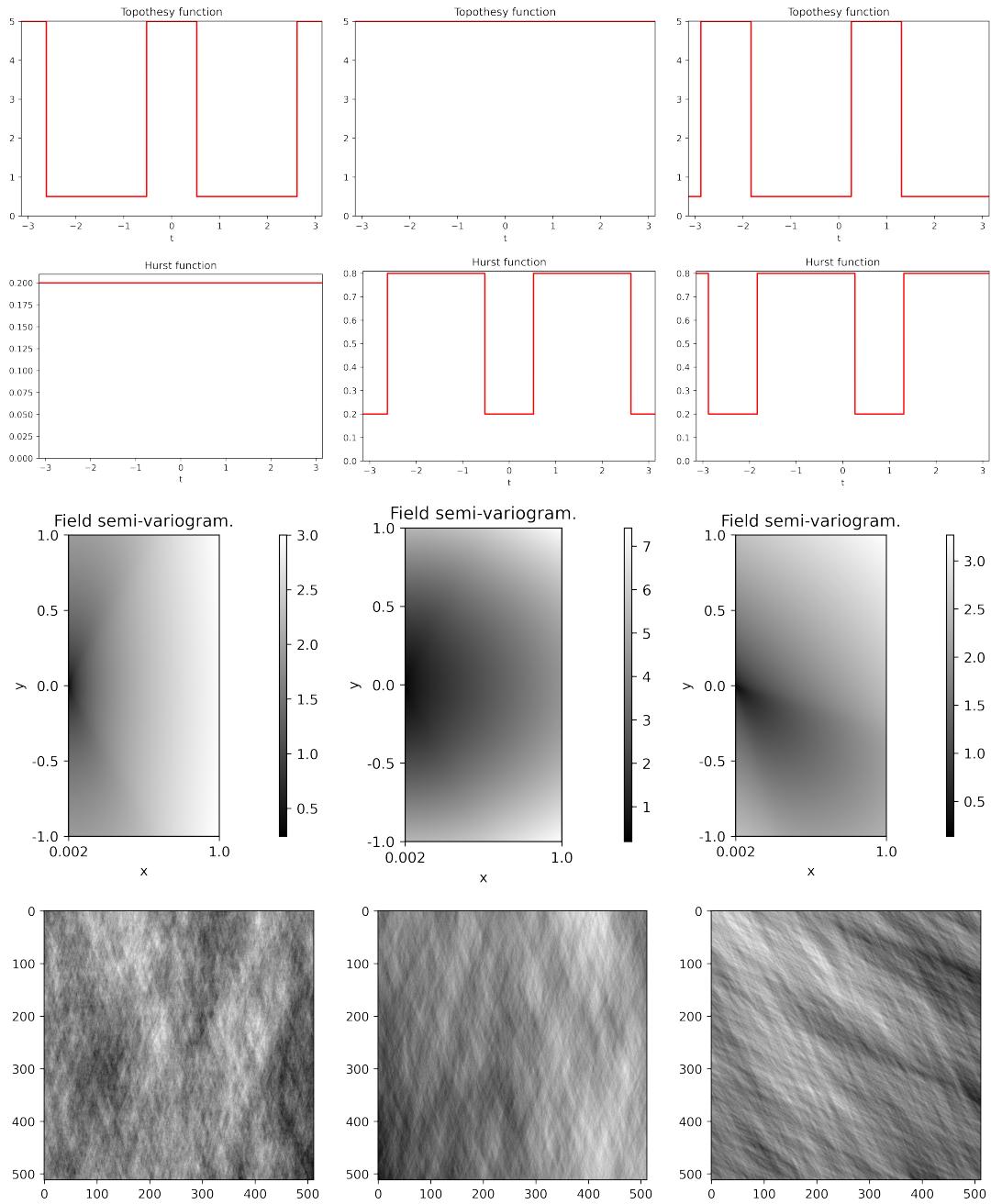


FIGURE 2.3 – Exemples de textures anisotropes avec τ et Hurst définies chacune par deux valeurs, respectivement, (0.5 et 5.0) et (0.2 et 0.8). Chaque colonne a une ouverture spécifique (0 ou $\frac{\pi}{3}$) pour τ et/ou Hurst, ainsi qu'une phase (respectivement 0,0 et $\frac{\pi}{4}$).

Chapitre 3

Analyse de textures homogènes par ondelettes

3.1 Formalisme et analyse multi-échelle

On va ici rapidement énoncer les définitions élémentaires du formalisme multi-échelle issu de la théorie des fractales. On verra ensuite que dans ce contexte, l'analyse par ondelettes, à savoir par translation et dilatation d'une famille de fonctions, est naturelle à considérer dans ce contexte.

De façon générale, on parle de formalisme fractal sur un champ X quand on a une quantité locale vérifiant

$$T_X(a, k) \sim \sigma_k a^{h_k} \quad \text{quand } a \rightarrow 0 \quad (3.1)$$

qui dépend de façon jointe de l'échelle a et de la position k . On peut associer à cette quantité une fonction de structure

$$S(a, q) = \frac{1}{n_a} \sum_{k=1}^{n_a} |T_X(a, k)|^q \sim a^{\zeta(q)}. \quad (3.2)$$

La fonction ζ est appelée fonction d'échelle et est obtenue par moyennage spatial à échelle a fixée des coefficients $|T_X(a, \cdot)|^q$.

D'un point de vue technique, on définit généralement ζ à partir du comportement de la fonction de structure aux plus petites échelles

$$\zeta(q) = \liminf_{j \rightarrow -\infty} \frac{\log(S(j, q))}{\log(2^j)}. \quad (3.3)$$

Dans notre contexte d'étude, l'objectif est de récupérer une information de régularité à partir de $\zeta(q)$. En fait, les champs considérés vérifiant une forme d'autosimilarité aux plus petites échelles, il est possible de montrer que pour de nombreuses quantités multi-échelles, on peut récupérer le coefficient d'autosimilarité à partir de ζ . On peut alors justifier l'application du lemme de Kolmogorov-Chentsov pour en déduire la régularité au sens de Hölder. On ne développe pas cette approche, qui permet également un lien avec la dimension de Hausdorff, par soucis de place.

Mentionnons l'article [Jaf+16] qui étudie en toute généralité le comportement de ζ en fonction de q et identifie $\zeta(q = \infty)$ comme régularité (critique) de Hölder de X .

3.2 Ondelettes pour l'analyse multi-échelle

Une façon naturelle d'obtenir des coefficients qui dépendent des échelles et des positions peut se faire à l'aide d'une analyse en ondelettes [Dav21]. Dans ce cadre là, pour des fonctions dans $L^2(\mathbb{R}^2)$, on va se donner une famille de fonctions ψ_1, \dots, ψ_D , appelées ondelettes génératrices, à partir desquelles on déduit les fonctions

$$\psi_{j,k}^{(d)} = 2^{-j} \psi^{(d)}(2^{-j} \cdot -k). \quad (3.4)$$

Ces fonctions, qui correspondent à des versions translatées et dilatées des ondelettes génératrices, permettent alors une analyse multi-échelle localisée de signaux X à partir des coefficients d'ondelettes

$$c_{j,k}^{(d)} = \langle X, \psi_{j,k}^{(d)} \rangle_{L^2}. \quad (3.5)$$

La littérature très développée sur les ondelettes, permet alors de donner des critères induisant certaines bonnes propriétés des coefficients d'ondelettes $c_{j,k}^{(d)}$. Ici, les propriétés d'intérêt vont être

1. La famille $\{\psi_{j,k}^{(d)}\}$ engendre $L^2(\mathbb{R}^2)$.
2. Les ondelettes ont suffisamment de moments nuls.
3. Les ondelettes sont à support compact.
4. Les coefficients sont numériquement stables.
5. La sélectivité directionnelle des ondelettes.

Justifions rapidement les raisons derrière ces choix. Tout d'abord la première propriété est très commune, généralement on demande même aux ondelettes d'être orthonormales. Cela permet d'être certain de ne pas être ‘aveugle’ à certaines composantes d'un signal.

La deuxième permet de garantir la capacité à identifier la régularité de Hölder à partir des coefficients [Jaf+16]. La propriété des (m -)moments nuls étant que pour tout polynôme P de degré inférieur ou égal à m , à coefficients (α_p) on a

$$\langle \psi, P \rangle = \sum_{p=0}^m \alpha_p \int t^p \psi(t) dt = 0. \quad (3.6)$$

Ainsi, l'analyse réalisée est orthogonale aux polynômes. D'un point de vue régularité on peut interpréter cela comme le fait que si un signal X est différentiable m -fois, alors on peut analyser la régularité de Hölder de $X^{(m)}$ le signal différentié m -fois. De même, on obtient les mêmes coefficients pour X ou $X + P$ avec P un polynôme de degré suffisamment petit. C'est à dire qu'on ignore les composantes polynomiales ou lisses du signal, ce qui correspond effectivement au point de vue analyse de rugosité qui est adopté ici.

L'importance du support compact, qui était l'une des principales motivations du développement initial des ondelettes, est que cela permet une analyse locale. C'est donc une propriété très désirable dans ce contexte de segmentation où l'on souhaite pouvoir déterminer localement la rugosité des textures. On verra en fait dans la partie suivante qu'un support compact fréquentiel est une propriété quasiment aussi importante qu'un support compact.

La stabilité numérique des coefficients, surtout par rapport aux translations des signaux analysés, est une propriété clef dans notre situation. En effet, on cherche à réaliser une analyse

aussi locale que possible, il est donc important que la valeur obtenue des coefficients dépende un minimum de la discréétisation réalisée. En effet, les ondelettes génératrices étant toujours des fonctions à moyenne nulle, les coefficients vont typiquement osciller fortement entre des valeurs négatives et positives, et donc même possiblement passer par 0. Ce genre de situation est classique en traitement du signal, une façon de faire est de passer par un signal dit analytique afin de déterminer l'enveloppe du signal. Le signal analytique X_a associé à un signal X s'obtient en dimension 1 sous la forme

$$X_a(t) = X(t) + i(HX)(t) \quad (3.7)$$

avec $(HX)(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{X(s)}{t-s} ds$ la transformée de Hilbert en t de X . On obtient ainsi X_a un signal à valeurs complexes dont le module correspond à la notion intuitive d'enveloppe. On peut mentionner que la phase de X_a correspond à une notion de fréquence instantanée.

Finalement, étant donné que l'objectif est de faire une analyse révélant des propriétés d'anisotropie, il est important que les ondelettes choisies soient au moins anisotropes. Par contre, une propriété désirable peut être que l'analyse de signaux isotropes, résulte dans des coefficients indépendants (en moyenne) de l'indice directionnel d . En se rappelant que les modèles de champs anisotropes X^f (cf Définition 2.3.4) considérés sont définis à partir de propriétés de f dans l'espace des fréquences, il est raisonnable de souhaiter pouvoir spécifier les propriétés directionnelles des ondelettes à partir d'une définition fréquentielle.

Afin de satisfaire à l'ensemble de ces contraintes, on utilise une transformée en ondelettes complexes due à Kingsbury, appelée Dual-Tree Complex Wavelet Transform (DTCWT) [SBK05]. Ces ondelettes (Cf. fig 3.1) sont construites par analogie avec la transformée de Hilbert mais réussissent à éviter de nombreuses mauvaises propriétés qui pourraient apparaître. De plus, mentionnons que l'analyse est particulièrement rapide en raison de sa construction qui est basée sur la mise en parallèle de deux algorithmes en banque de filtres de Mallat sur des ondelettes réelles [Mal89]. Mentionnons aussi qu'il est possible réaliser la synthèse du signal à partir des coefficients de DTCWT en parcourant ces banques de filtres en sens inverse.

Précisons que la version de l'algorithme utilisée est basée sur l'article [Kin06] qui permet une analyse isotrope des signaux isotropes¹. D'ailleurs cette version de la DTCWT a été introduite par Kingsbury dans un contexte très similaire au notre, à savoir qu'il souhaitait faire de la classification de champs élémentaires (Champs gaussiens autosimilaires anisotropes). Il a ainsi pu montrer que la DTCWT avait un meilleur comportement pour de la classification de textures anisotropes que l'ensemble des ondelettes traditionnelles telles que les ondelettes de Daubechies, de Coifman (coiflets), symlets et de Gabor.

3.3 Application sur AFBF et modèles

Ainsi, à partir de la DTCWT, on obtient pour chaque échelle j et position k , D coefficients d'ondelettes complexes $c_{j,k}^{(d)}$, avec $D = 6$ qui est interprété comme le nombre de directions disponibles. D'après la partie analyse multi-échelle, sur les AFBF X^f (cf. Définitions 2.3.3 et 2.3.4), chacune des suites de coefficients à d fixé, va suivre une loi de décroissance à travers les échelles en puissance caractérisée par un exposant $h_d \in]0, 1[$.

1. On peut par exemple constater que c'est le cas sur les figures 3.2 et 3.3 où la première image est isotrope, et les coefficients multi-échelles à travers les directions sont superposés.

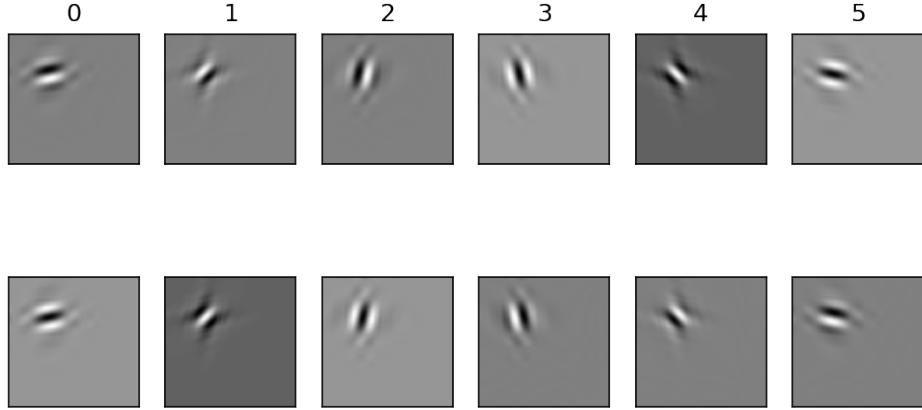


FIGURE 3.1 – Ondelettes complexes de [Kin06] direction par direction avec sur la première ligne la partie réelle et en dessous la partie imaginaire.

On peut alors utiliser l’expression dans le domaine spectral de l’analyse qui nous donne

$$\mathbb{E}|c_{j,k}^{(d)}|^2 = \mathbb{E}|\langle X^f, \psi_{j,k}^{(d)} \rangle|^2 = |\widehat{\langle f, \psi_{j,k}^{(d)} \rangle}|^2. \quad (3.8)$$

La puissance spectrale de X^f étant donnée par $f(\xi) = \tau(\arg \xi) ||\xi||^{-(2h(\arg \xi)+2)}$ et le support fréquentiel de $\psi_{j_1,k}^{(d)}$ à travers les échelles étant obtenu par multiplication du support de $\psi_{j_2,k}^{(d)}$ on déduit que $(\mathbb{E}|c_{j,k}^{(d)}|^2)_j$ décrit le comportement de f sur un cône fréquentiel défini par $\psi^{(d)}$. C’est pour cela que le support spectral compact des ondelettes est au moins aussi important que le support spatial compact.

Il est alors normal de s’intéresser aux log-coefficients $L_{j,k}^{(d)} = \log |c_{j,k}^{(d)}|^2$ qui vont donc suivre un modèle linéaire à travers les échelles. Ainsi, on a trois cas différents en fonction des propriétés de X^f , et donc par le théorème de Bonami-Estrade 2.3.2 de façon équivalente par les propriétés de f .

Modèles. En notant $L_{j,k}^{(d)}(X^f) = L_{j,k}^{(d)} = \log |\langle X^f, \psi_{j,k}^{(d)} \rangle|^2$ avec d ’après la Définition 2.3.3

$$f(\xi) = \tau(\arg \xi) ||\xi||^{-(2h(\arg \xi)+2)}.$$

On a les trois modèles suivants :

Si X^f est autosimilaire isotrope (τ et h constantes) :

$$L_{j,k}^{(d)} \sim jh + v. \quad (\text{isotrope})$$

Si X^f est autosimilaire anisotrope (τ non constante et h constante) :

$$L_{j,k}^{(d)} \sim jh + v_d. \quad (h\text{-lié})$$

Sinon :

$$L_{j,k}^{(d)} \sim j h_d + v_d. \quad (h\text{-libre})$$

C'est le modèle isotrope qui a été considéré pour de la segmentation de textures par Barbara Pascal dans [PPA21][Pas20]. Ici, les modèles auxquels nous nous intéressons sont les modèles anisotropes h -lié et h -libre. Afin de simplifier les notations pour la suite, on peut remarquer que ces modèles peuvent s'écrire

$$L_{j,k} \sim jH + V \quad (3.9)$$

en ayant noté $L_{j,k} = (L_{j,k}^{(d)})_{d=1}^D$ avec H et V dans \mathbb{R}^D . En contrignant les valeurs pouvant être prises par H et V on peut choisir le modèle étudié. En notant $1_D = (1, \dots, 1) \in \mathbb{R}^D$:

- isotrope : $H = h1_D$ et $V = v1_D$, avec $h, v \in \mathbb{R}$.
- h -lié : $H = h1_D$ et $V = (v_1, \dots, v_D)$, avec $h, v_1, \dots, v_D \in \mathbb{R}$.
- h -libre : $H = (h_1, \dots, h_D)$ et $V = (v_1, \dots, v_D)$, avec $h_1, \dots, h_D, v_1, \dots, v_D \in \mathbb{R}$.

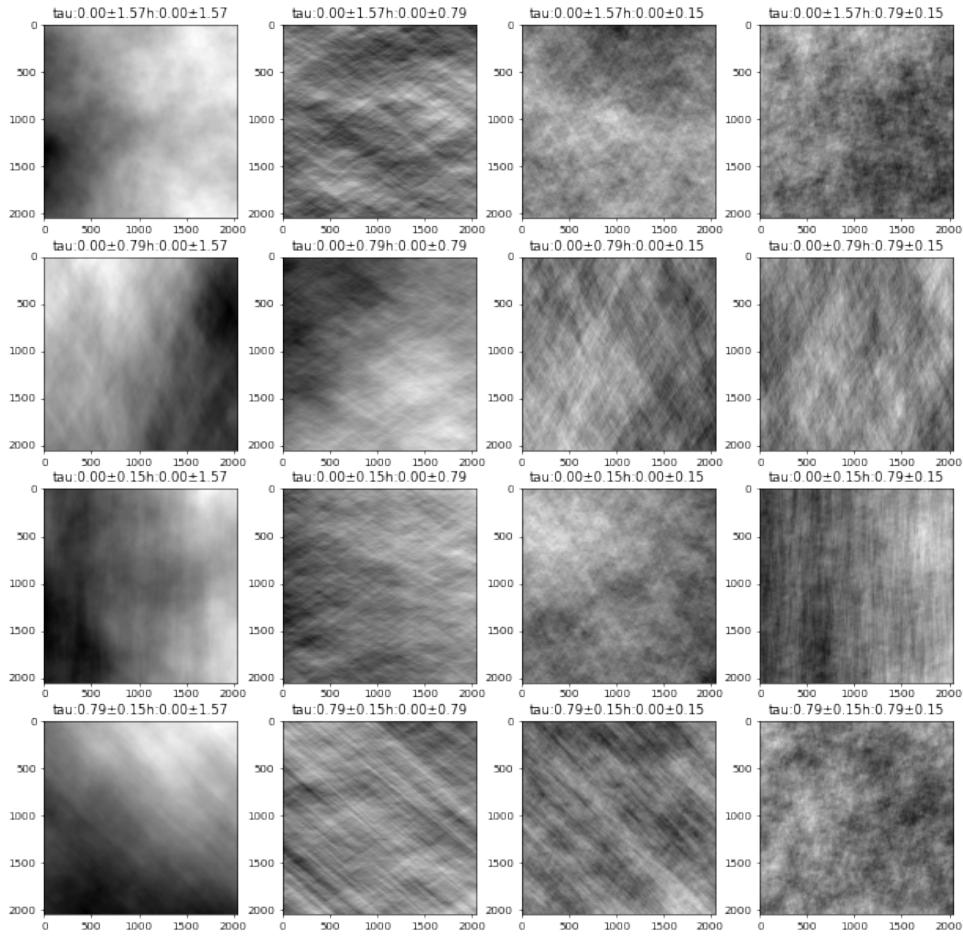


FIGURE 3.2 – Sur chacune des lignes (resp. colonnes) la fonction de topothèse τ (resp. de Hurst) est la même. De gauche à droite (resp. de haut en bas), la fonction de Hurst (resp. τ) a pour ouverture $\pi; \frac{\pi}{2}; 0.3; 0.3$, est centrée en $0; 0; 0; \frac{\pi}{4}$, et prends les valeurs 0.2 et 0.8 (resp. 0.5 et 5.0).

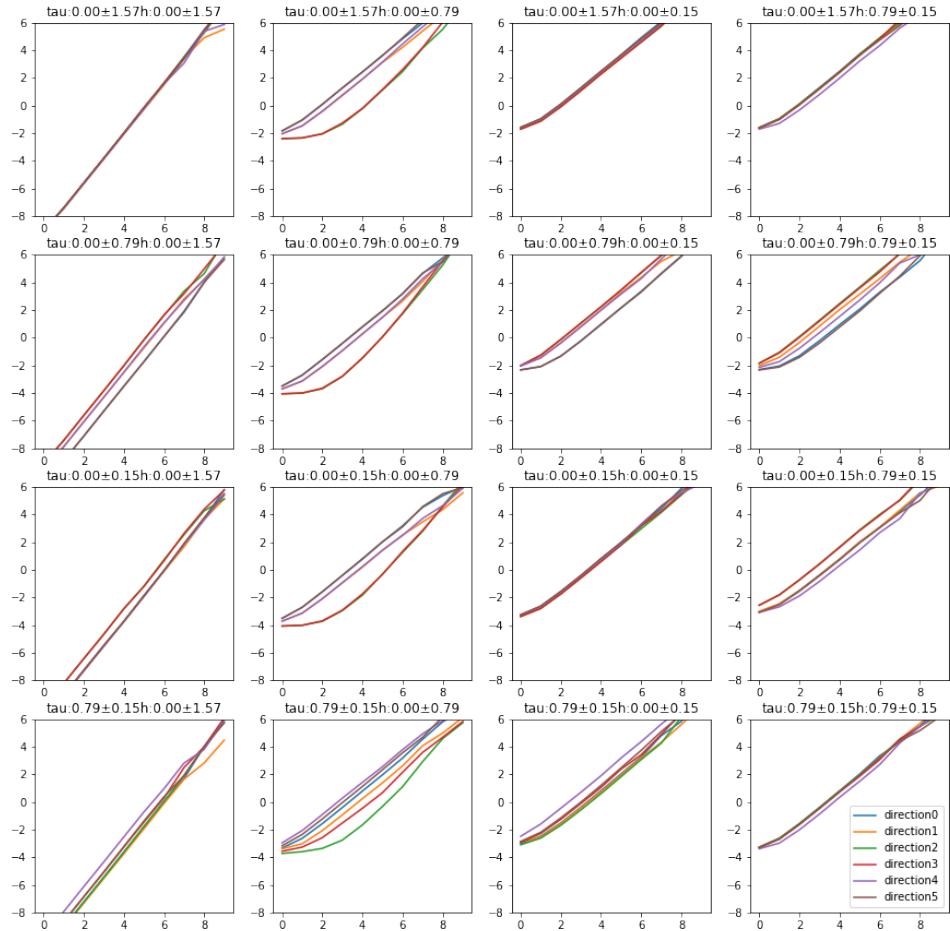


FIGURE 3.3 – log-coefficients d'ondelettes moyens des 16 textures de la figure 3.2 au travers des échelles. On constate que l'analyse permet de distinguer la quasi-totalité des textures présentées. Cependant, on constate aussi que ce modèle ne permet pas de distinguer n'importe quelle paire d'AFBF.

Chapitre 4

Optimisation et segmentation de textures hétérogènes

4.1 Estimation et segmentation

4.1.1 Stratégie générale

Les modèles autosimilaires homogènes isotrope, h -lié et h -libre présentés à la section précédente vérifient un modèle linéaire de la forme

$$L_{j,k}(X) = jH_k + V_k$$

où $L(X) = (\log(|\langle X, \psi_{j,k}^{(d)} \rangle|^2))_{d,j,k}$ sont les logs-coefficients des ondelettes $(\psi^{(d)})_{d=1}^D$ associés à X . De plus $H_k \in (0, 1)^D$ correspond à la régularité et dépend de la fonction de Hurst h et par ailleurs $V_k \in \mathbb{R}^D$ correspond plutôt à des propriétés induites par la fonction de topothèse τ .

Rappelons que la propriété de stationnarité de X permet d'estimer $jH+V$ par moyennage spatial tandis que la propriété d'autosimilarité permet de retrouver H et V au voisinage d'une certaine position k seulement à partir de $\{L_{j,k}(X) : j \leq J\}$ qui correspondent au coefficients d'ondelettes aux échelles les plus fines.

Bien que l'on cherche à exploiter ces deux propriétés deux difficultés se présentent. Tout d'abord le nombre d'échelles observables en pratique est fini, ce qui rend l'estimation ponctuelle des descripteurs imprécise. Ensuite, de façon plus spécifique à la segmentation de textures, le moyennage global sur l'image ne peut pas être utilisé (la texture étant hétérogène).

Afin de pouvoir résoudre le problème de segmentation une hypothèse est que les textures sont localement homogènes. Ainsi, le problème de sélection en chaque point d'un modèle linéaire défini à travers les échelles est contraint à être localement constant (à travers les positions). Cela va nous amener à imposer une pénalisation en norme 1 sur le gradient des solutions dans le but de favoriser les fonctions constantes par morceaux.

Ainsi, le modèle que l'on propose assure un compromis entre le modèle multi-échelle en chaque point et le fait qu'il y ait peu de variations de modèle sur des points proches. De plus, précisons que l'analyse que l'on réalise permet de détecter des propriétés d'anisotropie car on réalise une analyse à l'aide de D ondelettes, chacune ayant des propriétés d'orientation distinctes et toutes sont prises en compte dans l'analyse (cf. partie 3.2). Ceci est à contraster par exemple avec les travaux de Barbara Pascal [Pas20] qui reposent sur la définition des

coefficients d'ondelettes dominants[Jaf+16] qui sont des variables à une dimension obtenus par un maximum sur l'ensemble des directions, ce qui supprime l'information directionnelle.

Mentionnons également que traditionnellement le processus de segmentation pour des textures anisotropes était fait à part du processus d'estimation (estimation en deux temps). On peut par exemple citer l'article de Frédéric Richard [Ric16] qui réalise une segmentation basée sur la détection d'anisotropie¹. Dans cet article, un test statistique local est réalisé à valeur dans $\{0, 1\}$ qui détermine entièrement la segmentation. Le processus de segmentation se fait donc en deux étapes successives où la segmentation proposée n'influe pas sur le choix du test statistique local.

Ici, les deux processus sont joints, c'est à dire que l'on va itérer un compromis entre la géométrie des textures (localement constantes) et un modèle local multi-échelle des textures (log-log-linéaire). C'est à dire que le test statistique local prends en compte à la fois la qualité de la description locale de la texture, mais aussi la pertinence de la segmentation qui en résulterait (du point de vue de l'homogénéité locale).

Par exemple, une conséquence de ce modèle est une forte résistance aux *outliers* qui sont isolés dans une estimation homogène. On peut faire une analogie avec un modèle d'Ising où il est coûteux en énergie pour le système (et donc peu probable) d'avoir un spin entouré de spins de signes opposés (en tout cas à température suffisamment basse).

4.1.2 Attache(s) aux données et pénalisation

Notre procédure d'estimation repose sur la définition d'une fonction à minimiser de la forme $E + P$ avec

1. Un terme d'attache aux données E
2. Un terme de pénalisation P , permettant d'incorporer de l'information sur la solution attendue.

Le terme d'attache aux données dépend du modèle *h-libre* ou *h-lié* choisi dans la section précédente. Nous avons donc les deux choix anisotropes :

$$L_{j,k}^{(d)} = jh + v_d \quad (\text{i.e. } h\text{-lié})$$

ou bien,

$$L_{j,k}^{(d)} = jh_d + v_d. \quad (\text{i.e. } h\text{-libre})$$

Le premier modèle correspond à un modèle classique d'analyse multi-échelle d'un signal anisotrope autosimilaire, tandis que l'autre correspond à un modèle multi-échelle général.

On peut alors écrire les termes d'attache aux données en fonction du modèle choisi :

$$E_{\text{lié}}(V, H) = E_{\text{lié}} = \frac{1}{2} \sum_j \sum_d \|v_d + jh - L_j^{(d)}\|_F^2 \quad (4.1)$$

$$E_{\text{libre}}(V, H) = E_{\text{libre}} = \frac{1}{2} \sum_j \sum_d \|v_d + jh_d - L_j^{(d)}\|_F^2. \quad (4.2)$$

1. Le test qui est fait dans[Ric16] est un test qui détermine localement si la texture est isotrope, ici on va au delà de cette étude car on cherche aussi à distinguer des textures ayant des propriétés d'anisotropie distinctes. Le lecteur intéressé pourra trouver une application en neurosciences d'images microscopiques ainsi qu'une évaluation statistique de la capacité à distinguer des AFBF selon l'intensité de l'anisotropie portés par la fonction τ ou la fonction de Hurst.

Rappelons que chacun des termes qui apparaissent dans les problèmes précédents sont des matrices, interprétées comme des vecteurs, de la taille de l'image originale.

Finalement, on peut noter A l'application linéaire

$$(V, H) \mapsto A(V, H) = (V + jH)_j \quad (4.3)$$

qui à un coefficient d'ordonnée à l'origine V et un coefficient de pente H (les deux quantités pouvant être vectorielles et dépendantes du choix du modèle), associe les points correspondants à travers les échelles j . On peut remarquer que les problèmes de minimisation des termes d'attache aux données sont des problèmes de régression linéaire au sens des moindres carrés et admettent des formes closes (explicitées en 4.5). En effet, en remarquant

$$F = \sum_j \sum_d \|v_d + jh_d - L_j^{(d)}\|_F^2 = \|A(V, H) - L_j\|_F^2. \quad (4.4)$$

puis, en différenciant et en calculant la valeur pour laquelle le gradient vaut 0 (critère de Fermat), on trouve la forme close classique

$$(V_{LR}, H_{LR}) = (A^* A)^{-1} A^*(L_j)_j \quad (4.5)$$

qui correspond aux coefficients de régression linéaire.

On peut voir fig. 4.1 des solutions de ces problèmes et on constate qu'il y a une forte variance au niveau des coefficients. Afin d'avoir des solutions constantes par morceaux et de pouvoir retrouver les ‘vrais’ coefficients il est nécessaire de régulariser les solutions. Cela passe par le choix d'un terme de pénalisation. Dans cet objectif, on va utiliser une généralisation de la

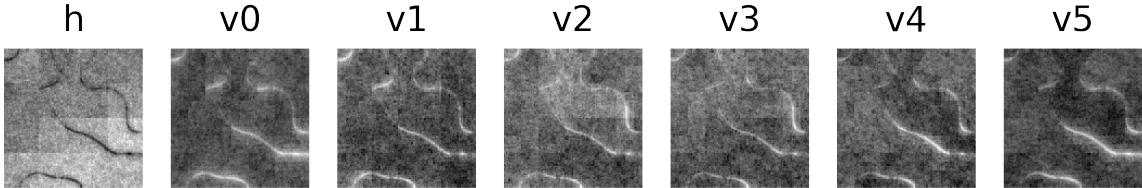


FIGURE 4.1 – Coefficients solutions de 4.1 pour deux textures homogènes anisotropes mélangées (cf. simulation 4.5.1).

variation totale qui est un terme de pénalisation classique pour obtenir des solutions constantes par morceaux. Pour cela, on va choisir une norme 1 appliquée à une norme euclidienne sur le gradient des coefficients estimés $\|\nabla(V, H)\|_{1,2} = \|(\|\nabla(V, H)\|_2)\|_1$ qui est une forme de généralisation de la variation totale en dimension > 1 (ici $D + 1$ ou $2D$).

Explicitement, on pose un opérateur de gradient finie horizontal (l'opérateur vertical G_y est défini de façon analogue)

$$(G_x u)_{(k)} = u_{(k+e_x)} - u_{(k)} \quad (4.6)$$

on définit alors un terme de pénalisation de la forme

$$\text{TV}_\alpha(V, H) = \sum_k \sqrt{\sum_d |G_x v_{d,k}|^2 + |G_y v_{d,k}|^2 + \alpha (|G_x h_{d,k}|^2 + |G_y h_{d,k}|^2)}.$$

Une remarque qui sera importante par la suite, est que cette variation totale peut s'écrire

$$\text{TV} = \|\cdot\|_{2,1} \circ G \quad (4.7)$$

avec G la concaténation des opérateurs de gradients horizontaux et verticaux.

Cette définition de variation totale nous permet à l'aide d'un hyperparamètre α d'appliquer des contraintes plus ou moins forte sur les v_d ou sur les h_d . Remarquons que cette pénalisation pour le modèle h -libre s'adapte au modèle h -lié en prenant $h_d = \frac{1}{\sqrt{D}}h$ pour chacune des orientations $d = 1, \dots, D$.

4.1.3 Modèles

Au vu des fonctions d'attache aux données et de pénalisation définies dans les sections précédentes on considère les deux fonctionnelles à minimiser :

$$\underset{V,H}{\text{minimiser}} : \frac{1}{2} \sum_j \sum_d \|v_d + jh_d - L_j^{(d)}\|_F^2 + \lambda \text{TV}_\alpha(V, H) \quad (h\text{-libre-min})$$

et

$$\underset{V,H}{\text{minimiser}} : \frac{1}{2} \sum_j \sum_d \|v_d + jh - L_j^{(d)}\|_F^2 + \lambda \text{TV}_\alpha(V, H) \quad (h\text{-lié-min})$$

Pour ces deux fonctionnelles le terme λ donne le niveau de régularisation à apporter aux solutions. Si $\lambda = 0$, aucune régularisation n'est appliquée et les solutions sont données par (4.5). De façon analogue, si λ est choisi arbitrairement grand, le terme d'attache aux données devient négligeable et la solution est constante sur toute l'image. Il faut donc réaliser un compromis sur le choix de λ afin d'avoir une solution constante par morceaux qui reste pertinente pour expliquer les données.

En optimisation, face à des problèmes de minimisation la stratégie la plus classique consiste à appliquer une méthode de descente de gradient. La difficulté pour toutes ces fonctionnelles provient du fait que TV n'est pas lisse ce qui empêche d'en calculer le gradient. On peut cependant remarquer que ces fonctionnelles appartiennent à $\Gamma_0(\mathcal{H})$ l'ensemble des fonctions réelles convexes semi-continues inférieurement et propres. Il se trouve que pour cette classe de problèmes il existe une classe de stratégies de minimisation analogue à une descente de gradient basée sur la théorie des opérateurs proximaux, ce qui sera le sujet de la prochaine section.

Avant de passer à la formulation générale des problèmes proximaux qui nous amènera à des algorithmes permettant la minimisation des fonctionnelles ci-dessus, énonçons le lemme suivant qui nous permet de remplacer les coefficients d'ondelettes directement par les coefficients du modèle linéaire associé. L'intérêt est que le terme d'attache aux données dépendra alors seulement des différences entre le modèle linéaire testé et le modèle linéaire optimal au sens des moindres carrés et dépend seulement de façon implicite des coefficients d'ondelettes.

Lemme 4.1.1. *En posant par définition les coefficients de régression linéaire,*

$$(V_{LR}, H_{LR}) = (A^* A)^{-1} A^* (L_j)_j \quad (4.8)$$

avec A défini en 4.3. Alors,

$$\sum_j \|V + jH - L_j\|_F^2 = C + \sum_j \|V + jH - V_{LR} - jH_{LR}\|_F^2 = C + \|A(V - V_{LR}, H - H_{LR})\|_F^2. \quad (4.9)$$

où C est une constante. De plus,

$$\|A(V - V_{LR}, H - H_{LR})\|_F^2 \quad (4.10)$$

est fortement convexe.

Démonstration. En effet,

$$\begin{aligned} \sum_j \|V + jH - L_j\|_F^2 &= \|(V + jH)_j - (L_j)_j\|_F^2 \\ &= \|A(V, H) - (L_j)_j\|_F^2 \\ &= \|A(V - V_{LR}, H - H_{LR}) + A(V_{LR}, H_{LR}) - (L_j)_j\|_F^2 \end{aligned}$$

où V_{LR} et H_{LR} sont les ordonnées à l'origine et les pentes des régressions linéaires, et sont donc données par

$$\begin{aligned} (V_{LR}, H_{LR}) &= (A^*A)^{-1}A^*(L_j)_j. \quad (4.11) \\ &= \|A(V - V_{LR}, H - H_{LR}) + A(V_{LR}, H_{LR}) - (L_j)_j\|_F^2 \\ &= \|A(V - V_{LR}, H - H_{LR})\|_F^2 + \|A(V_{LR}, H_{LR}) - (L_j)_j\|_F^2 \\ &\quad + 2\langle A(V - V_{LR}, H - H_{LR}), A(V_{LR}, H_{LR}) - (L_j)_j \rangle. \end{aligned}$$

En remarquant que

$$\begin{aligned} \langle A(V - V_{LR}, H - H_{LR}), A(V_{LR}, H_{LR}) - (L_j)_j \rangle \\ = \langle (V - V_{LR}, H - H_{LR}), A^*A(V_{LR}, H_{LR}) - A^*(L_j)_j \rangle \\ = \langle (V - V_{LR}, H - H_{LR}), A^*A(A^*A)^{-1}A^*(L_j)_j - A^*(L_j)_j \rangle = 0. \end{aligned}$$

On a alors que la constante qui apparaît est l'erreur quadratique d'approximation du meilleur modèle au sens des moindres carrés.

Concernant la forte convexité, il suffit par un calcul explicite de remarquer que A^*A est une matrice symétrique définie positive, la constante de forte convexité correspond alors à la valeur propre minimale de A^*A . \square

4.2 Opérateurs proximaux

4.2.1 Théorie générale et formulation duale

L'ensemble des fonctionnelles à minimiser proposées ci-dessus sont des problèmes convexes semi-continus inférieurement mais non lisses. Ainsi, on ne peut pas appliquer les approches de minimisation classiques par descente de gradient. Cependant, la définition de sous-différentielle de Moreau nous permettra d'étendre la notion de gradient aux cas non-différentiable. Plus précisément, en notant $\Gamma_0(\mathcal{H})$ l'ensemble des fonctions réelles convexes semi-continues inférieurement² sur \mathcal{H} à valeurs dans $\mathbb{R} \cup \{+\infty\}$ on peut énoncer la définition :

2. Une fonction est dite semi-continue inférieurement si et seulement si son épigraphe est fermé, en analyse convexe une telle fonction est souvent appelée fermée. Pour les fonctions convexes, l'épigraphe est alors fermé et convexe ce qui garantit l'existence d'au moins un minimiseur. De plus, pour toute position initiale et minimiseur dans \mathcal{H} il existe alors un segment fermé reliant les deux sur lequel la fonction est monotone, convexe et atteint ses bornes.

Définition 4.2.1. A une fonction $\varphi \in \Gamma_0(\mathcal{H})$, on peut lui associer une sous-différentielle définie par

$$\partial\varphi : z \mapsto \{s \in \mathcal{H} : \forall w \in \mathcal{H}, \langle w - z, s \rangle \leq \varphi(w) - \varphi(z)\}. \quad (4.12)$$

Afin de résoudre ces problèmes nous avons besoin d'avoir recours à des approches proximales que nous allons décrire dans cette section et qui repose sur la notion d'opérateur proximal³. Nous pourrons ainsi considérer des algorithmes adaptés aux problèmes d'optimisation h -libre-min et h -lié-min.

Théorème 4.2.2. [BC+11] Soit \mathcal{H} un espace de Hilbert réel et $\|\cdot\|$ la norme induite par son produit scalaire,

- (Moreau) Soit $x \in \mathcal{H}$ et $\varphi \in \Gamma_0(\mathcal{H})$, alors le problème de minimisation

$$\underset{y \in \mathcal{H}}{\text{minimiser}} : \varphi(y) + \frac{1}{2}\|x - y\|^2 \quad (4.13)$$

admet une unique solution notée $\text{prox}_\varphi(x)$ appelée opérateur proximal de φ en x .

- (Hiriart-Urruty) Soit $\mathcal{H} = \mathbb{R}^N$ et soit U une matrice symétrique définie positive de taille $N \times N$ et $\|\cdot\|_U = \langle \cdot, U \cdot \rangle$ la métrique induite par U sur \mathcal{H} , alors l'opérateur proximal de φ dans la métrique induite par U est défini comme l'unique solution de

$$\underset{y \in \mathcal{H}}{\text{minimiser}} : \varphi(y) + \frac{1}{2}\|x - y\|_U^2. \quad (4.14)$$

et est noté prox_φ^U .

L'un des intérêts, notamment du point de vue des applications, est que pour de nombreuses fonctions il existe des formes closes du prox correspondant. Par exemple, si on considère $\varphi = \|\cdot\|_1$, alors le prox correspondant est un seuillage doux

$$\text{prox}_{\tau\|\cdot\|_1} : x \mapsto \text{sign}(x) \max(0, |x| - \tau). \quad (4.15)$$

Il se trouve que les termes de pénalisation et de régularisation qui sont considérés ici sont tous dans $\Gamma_0(\mathcal{H})$ et on peut donc y appliquer les méthodes proximales. Par exemple, les problèmes de minimisation que nous considérons admettent chacun une unique solution qui peut s'écrire de façon concise sous la forme de Hiriart-Urruty :

$$\text{prox}_{\lambda \text{TV}_\alpha}^{A^*A}(V_{LR}, H_{LR}) := \arg \min_{V, H} \frac{1}{2} \|(V - V_{LR}, H - H_{LR})\|_{A^*A}^2 + \lambda \text{TV}_\alpha(V, H). \quad (4.16)$$

Une difficulté de cette formulation est qu'il n'existe pas de forme close donnant la solution. Les algorithmes proposés reposeront sur d'une part sur une notion de dualité convexe et d'autre part feront appel à des algorithmes de splitting (entre problème primal et dual).

Introduisons maintenant la définition de la fonction conjuguée dont nous aurons besoin afin de pouvoir ensuite exprimer le lien entre problème primal et problème dual.

Définition 4.2.3. A une fonction $\varphi \in \Gamma_0(\mathcal{H})$ on peut associer la fonction conjuguée convexe :

$$\varphi^*(z) = \sup_w \langle z, w \rangle - \varphi(w) \quad (4.17)$$

3. On peut citer le site <http://proximity-operator.net/> qui regroupe un grand nombre de résultats et références utiles pour l'utilisation des opérateurs proximaux.

Ainsi, en remarquant que les problèmes d'optimisations considérés peuvent s'écrire sous la forme

$$\text{minimiser} : E(z) + P(Lz) \quad (4.18)$$

avec E qui correspond à $E_{\text{lié}}$ ou E_{libre} (cf 4.1 et 4.2), G qui correspond à $\|\cdot\|_{2,1}$ et L qui correspond à G (cf 4.7), on va pouvoir appliquer le théorème suivant pour obtenir à la fois le problème dual et la relation entre solution primale et duale.

Théorème 4.2.4 (Théorème 19.1 [BC+11]). *Au problème*

$$\text{minimiser} : E(z) + P(Lz) \quad (4.19)$$

avec $E \in \Gamma_0(\mathcal{H})$, $P \in \Gamma_0(\mathcal{G})$ et L un opérateur linéaire continu de \mathcal{H} dans \mathcal{G} , on peut associer le problème dual

$$\text{minimiser}_u : E^*(-L^*u) + P^*(u). \quad (4.20)$$

De plus, si il existe \hat{u} vérifiant,

$$0 \in \partial P^*(\hat{u}) - L \partial E^*(-L^*\hat{u}) \quad (4.21)$$

alors \hat{u} est une solution du problème dual. Finalement, il existe une solution \hat{z} du problème primal vérifiant

$$L\hat{z} \in \partial P^*(\hat{u}) \quad \text{et} \quad -L^*\hat{u} \in \partial E(\hat{z}). \quad (4.22)$$

Dans notre cas on a donc le problème dual

$$\text{minimiser}_{(s,w)} : F^*(-D^*(s,w)) + \|(s,w)\|_{2,1}^* \quad (\text{dual})$$

en notant F^* la fonction conjuguée du terme d'attache aux données choisi (soit $E_{\text{lié}}$ soit E_{libre}) et (s,w) la variable duale à (V,H) . Ainsi, si on a (\hat{s}, \hat{w}) comme solution du problème dual, la solution du problème primal (\hat{V}, \hat{H}) est obtenue par

$$-D^*(\hat{s}, \hat{w}) = \partial F^*((\hat{V}, \hat{H})). \quad (4.23)$$

Finalement, nous aurons besoin de deux résultats classiques supplémentaires afin de pouvoir décrire les algorithmes. Tout d'abord, énonçons la relation de Moreau qui permet de relier l'opérateur proximal d'une fonction et de sa conjuguée :

Théorème 4.2.5. [BC+11] Soit $\phi \in \Gamma_0(\mathcal{H})$ et $\gamma > 0$, alors

$$z = \text{prox}_{\gamma\phi^*}(z) + \gamma \text{prox}_{\phi/\gamma}\left(\frac{z}{\gamma}\right). \quad (4.24)$$

Le dernier résultat dont nous aurons besoin et que nous ne démontrerons pas est l'expression, en forme close, du prox associé à $\|\cdot\|_{2,1}$:

Proposition 4.2.6. [Chi+17] Soit $z = (z_{i,j})_{i,j} \in \mathbb{R}^{M \times D}$, en notant

$$\|z\|_{2,1} = \sum_i^M \sqrt{\sum_j^D |z_{i,j}|^2} \quad (4.25)$$

on a

$$\text{prox}_{\gamma\|\cdot\|_{2,1}}(z) = \left(\left(1 - \frac{\gamma}{\max(\|z_i\|_{\ell^2(D)}, \gamma)}\right) z_i \right)_{1 \leq i \leq M}. \quad (4.26)$$

4.2.2 Algorithmes proximaux

Dans la théorie des opérateurs proximaux, de nombreux algorithmes sont disponibles. Les algorithmes étant itératifs, on va les définir sous forme schéma itératif sous la forme

$$z_{t+1} = \Phi(z_t). \quad (4.27)$$

Le choix de l'algorithme Φ à utiliser en pratique dépend fortement de la situation dans laquelle on se trouve, notamment des opérateurs proximaux pour lesquels on connaît des formes closes.

Énonçons quelques algorithmes, avec τ désignant le pas de l'algorithme, afin d'expliquer les choix qui ont été faits pour la résolution du problème primal (h -lié-min, h -libre-min) ou dual (dual)

$$\underset{z}{\text{minimiser}} : E(z) + P(z). \quad (4.28)$$

1. Descente de gradient :

$$\Phi = I - \tau \nabla E - \tau \nabla P \quad (4.29)$$

2. Point proximal :

$$\Phi = \text{prox}_{\tau(E+P)} \quad (4.30)$$

3. Forward-backward :

$$\Phi = \text{prox}_{\tau E}(I - \tau \nabla P) \quad (4.31)$$

4. Douglas-Rachford :

$$\Phi = \text{prox}_{\tau E}(2\text{prox}_{\tau P} - I) + I - \text{prox}_{\tau E} \quad (4.32)$$

On remarque ainsi que la descente de gradient n'est pas applicable ici car le terme de régularisation n'est pas lisse. Le point proximal pourrait être appliqué si on avait connaissance d'une forme close de son expression. Cependant, il est rare d'avoir une forme close de l'opérateur proximal pour une somme de deux fonctions [PC17].

Par contre, forward-backward peut être facilement appliquée avec P le terme (conjugué) d'attache aux données qui est différentiable. De même on pourrait appliquer Douglas-Rachford sous réserve du calcul du prox de $F^* \circ (-G^* \cdot)$ pour le problème dual, ou du prox de $\|\cdot\|_{2,1} \circ G$ pour le problème primal. Le choix a été fait ici d'implémenter forward-backward, notamment motivés par sa simplicité, ainsi que le fait qu'il ait déjà été appliquée dans un cas analogue par Barbara Pascal [PPA21]. Pour cet algorithme le théorème suivant nous garantit une convergence vers un minimiseur de la fonctionnelle d'intérêt :

Théorème 4.2.7 (Proposition 28.13 [BC+11]). *Soit $E, P \in \Gamma_0(\mathcal{H})$ avec E différentiable et un gradient K -Lipschitz tel que*

$$\underset{z}{\text{minimiser}} : E(z) + P(z) \quad (4.33)$$

admet une solution. Soit $z_0 \in \mathcal{H}$ et

$$z_{t+1} = \text{prox}_{\tau P}(z_t - \tau \nabla E(z_t)), \quad (4.34)$$

si $0 < \tau < \frac{2}{K}$, alors la suite $(z_t)_t$ converge vers un minimiseur de $E + P$.

De plus, si $\delta = \min_z E(z) + P(z)$, alors

$$|E(z_t) + P(z_t) - \delta| = \mathcal{O}\left(\frac{1}{t}\right). \quad (4.35)$$

4.2.3 Problème dual

L'objectif ici est d'expliciter la fonctionnelle duale au problème considéré ainsi que le terme conjugué d'attache aux données. On pourra alors appliquer le théorème 4.2.7.

Réécrivons le problème primal afin de pouvoir y appliquer le théorème 4.2.4 :

$$(\hat{V}, \hat{H}) = \operatorname{argmin} \frac{1}{2} \|A(V - V_{LR}, H - H_{LR})\|_F^2 + \lambda \|G(V, H)\|_{2,1} \quad (4.36)$$

Il nous faut donc ensuite calculer la fonction conjuguée convexe de notre terme d'attache aux données :

$$F_{V_{LR}, H_{LR}}(V, H) := F(V, H) := \frac{1}{2} \|A(V - V_{LR}, H - H_{LR})\|_F^2. \quad (4.37)$$

Lemme 4.2.8. Avec F défini ci-dessus et (V_{LR}, H_{LR}) défini en 4.5 :

$$F^*(V, H) = -\frac{1}{2} \langle (V, H), (A^* A)^{-1}(V, H) \rangle + \langle (V_{LR}, H_{LR}), (V, H) \rangle. \quad (4.38)$$

De plus F^* est différentiable et son gradient s'écrit :

$$\nabla F^*(V, H) = -(A^* A)^{-1}(V, H) + (V_{LR}, H_{LR}). \quad (4.39)$$

Démonstration. Par définition la conjuguée convexe F^* de F correspond à

$$\begin{aligned} -F^*(V, H) &= \inf_{t,u} F(t, u) - \langle (t, u), (V, H) \rangle \\ &= \inf_{t,u} \frac{1}{2} \|A(t - V_{LR}, u - H_{LR})\|_F^2 - \langle (t, u), (V, H) \rangle \\ &= \inf_{t,u} \frac{1}{2} \langle (t - V_{LR}, u - H_{LR}), A^* A(t - V_{LR}, u - H_{LR}) \rangle - \langle (t, u), (V, H) \rangle \end{aligned}$$

La fonction à minimiser étant différentiable, on applique le critère de Fermat pour calculer la valeur minimale en cherchant pour quelle variable (t, u) la différentielle est nulle.

$$A^* A(t - V_{LR}, u - H_{LR}) - (V, H) =^{(hyp.)} 0 \quad (4.40)$$

et ainsi

$$\begin{aligned} -F^*(V, H) &= \frac{1}{2} \langle (A^* A)^{-1}(V, H), (V, H) \rangle - \langle (A^* A)^{-1}(V, H) + (V_{LR}, H_{LR}), (V, H) \rangle \\ &= \frac{1}{2} \langle (V, H), (A^* A)^{-1}(V, H) \rangle - \langle (V_{LR}, H_{LR}), (V, H) \rangle \end{aligned}$$

et on a le gradient de F^*

$$\nabla F^*(V, H) = -(A^* A)^{-1}(V, H) + (V_{LR}, H_{LR}). \quad (4.41)$$

□

A l'aide du Lemme 4.2.8 on peut expliciter le problème dual :

$$\underset{(s,w)}{\text{minimiser}} : \frac{1}{2} \langle G^*(s, w), (A^* A)^{-1}G^*(s, w) \rangle - \langle (V_{LR}, H_{LR}), G^*(s, w) \rangle + \|(s, w)\|_{2,1}. \quad (4.42)$$

On peut aussi calculer explicitement la relation entre solution primaire et duale en remplaçant le calcul explicite du gradient de F^* dans 4.23. En notant (\hat{s}, \hat{w}) une solution du problème dual, on a comme solution du problème primal :

$$(A^* A)((V_{LR}, H_{LR}) - G^*(\hat{s}, \hat{w})) = (\hat{V}, \hat{H}). \quad (4.43)$$

4.3 Forward-Backward et version accélérée

On peut à présent écrire un premier algorithme de minimisation sur le problème dual

$$\text{minimiser} : F^*(-G^*(s, w)) + \|(s, w)\|_{2,1}^* \quad (4.44)$$

en appliquant l'algorithme Forward-Backward. C'est à dire en appliquant successivement une minimisation par descente de gradient par rapport à $F^* \circ (-G \cdot)$ suivi d'un opérateur proximal de $\|\cdot\|_{2,1}^*$ appliqué en la nouvelle position. A partir du Théorème 4.2.7, nous pouvons écrire les itérations suivantes :

$$(s_{t+1}, w_{t+1}) = \text{prox}_{\tau \|\cdot\|_{2,1}^*}((s_t, w_t) + \tau G(-(A^* A)^{-1} G^*(s_t, w_t) + (V_{LR}, H_{LR}))$$

Il reste à calculer la constante de Lipschitz du gradient de $F^* \circ G^*$ qui est majorée par

$$\lambda_{\max}((A^* A)^{-1}) \lambda_{\max}(GG^*) = \frac{\lambda_{\max}(GG^*)}{\lambda_{\min}(A^* A)} \quad (4.45)$$

en utilisant que le rayon spectral coïncide avec la norme d'opérateur pour les matrices qui s'écrivent $Z^* Z$.

On peut alors conclure d'après le théorème 4.2.7 que les itérés convergent vers une solution du problème dual si $\tau < \frac{2\lambda_{\min}(A^* A)}{\lambda_{\max}(GG^*)}$. En appliquant la relation de Moreau (Théorème 4.2.5), on peut réécrire le schéma itératif sous la forme

$$\begin{cases} r_t &= G(-(A^* A)^{-1} G^*(s_t, w_t) + (V_{LR}, H_{LR})) \\ (s_t, w_t) &= (s_t, w_t) + \tau r_t - \tau \text{prox}_{\frac{1}{\tau} \|\cdot\|_{2,1}}(\frac{1}{\tau}(s_t, w_t) + r_t). \end{cases}$$

Pour cet algorithme, on a que la distance à la valeur minimale du problème dual décroît en $\mathcal{O}(\frac{1}{t})$. Il est important de noter que ce ne sont pas les itérés qui convergent à ce rythme vers la bonne solution.

En considérant que prox est simplement un seuillage doux, à chaque pas on met à jour la solution seulement dans les directions où il y a une grande différence avec les différences finies sur les solutions de la régression linéaire. La contractivité au voisinage de 0 de prox force les coefficients suffisamment petits à tendre vers 0 à travers les itérations, et ainsi les solutions à être constantes par morceaux.

On peut aussi remarquer que cet algorithme de minimisation du problème dual correspond à un algorithme ISTA (*Iterative Soft Thresholding Algorithm*). On peut alors écrire l'algorithme sous la forme accélérée FISTA (*Fast ISTA*) en prenant soin d'expliciter la variable primale (d'intérêt) $z = (V, H)$ par rapport aux variables duals u, \tilde{u} :

$$\begin{cases} u_{t+1} &= \text{prox}_{\tau \|\cdot\|_{2,1}^*}(\tilde{u}_t + \tau_t G z_t) \\ \tau_{t+1} &= \frac{t+b-1}{b} \\ \tilde{u}_{t+1} &= u_{t+1} + \frac{\tau_t-1}{\tau_{t+1}}(u_{t+1} - u_t) \\ z_{t+1} &= z_t - (A^* A)^{-1} G^*(u_{t+1} - u_t). \end{cases} \quad (4.46)$$

Pour cet algorithme, le théorème suivant a été démontré par Chambolle et Dossal

Théorème 4.3.1. *Corollaire 2[CD15] Avec (u_t) la suite définie par FISTA 4.46 avec $b > 2$, alors en notant u^* le minimiseur du problème dual, on a*

$$\|u_t - u^*\| = \mathcal{O}\left(\frac{1}{t^2}\right) \quad (4.47)$$

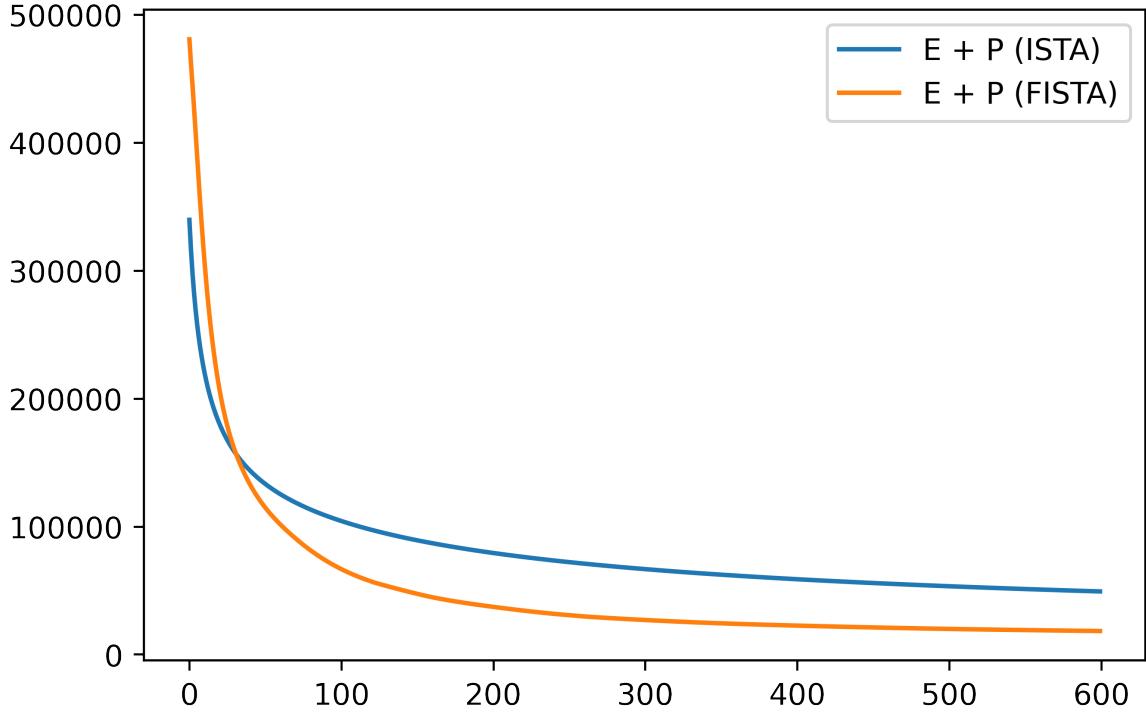


FIGURE 4.2 – Comparaison de la vitesse de convergence entre ISTA et FISTA

4.4 Algorithme primal dual et accélération par forte convexité

Une autre alternative repose sur l'algorithme de Chambolle-Pock [CP11] qui utilise explicitement la relation avec le problème primal pour résoudre un problème de la forme 4.19. L'algorithme primal-dual considéré s'écrit de la façon suivante pour un terme d'attache aux données E , un terme de pénalisation P et un opérateur linéaire L :

$$\begin{cases} z_{t+1} = \text{prox}_{\sigma E}(z_t - \sigma L^* \tilde{u}_t) \\ u_{t+1} = \text{prox}_{\tau P^*}(u_t + \tau L z_{t+1}) \\ \tilde{u}_{t+1} = u_{t+1} + \theta(u_{t+1} - u_t). \end{cases}$$

La formulation originale de cet algorithme est pour la résolution du problème de point-selle

$$\min_z \max_u \langle Lz, u \rangle + E(z) - P^*(u) \quad (4.48)$$

qui correspond à la formulation primal-dual du problème dual

$$\text{minimiser : } E^*(-L^*u) + P^*(u). \quad (4.49)$$

et du problème primal

$$\text{minimiser : } E(z) + P(Lz). \quad (4.50)$$

Ainsi, l'objectif de l'algorithme de Chambolle-Pock est de calculer une paire de points (\hat{z}, \hat{u}) qui minimisent simultanément le problème primal et le problème dual. C'est une telle paire de points qui s'appelle un point selle.

Afin de pouvoir appliquer cet algorithme nous avons donc besoin de calculer le prox de l'attache aux données F .

Lemme 4.4.1.

$$\text{prox}_{\gamma F} = (I + \gamma A^* A)^{-1}(\gamma A^* A(V_{LR}, H_{LR}) + (V, H)) \quad (4.51)$$

Démonstration. Par définition,

$$(\hat{V}, \hat{H}) = \text{prox}_{\gamma F}(V, H) = \arg \min_{V', H'} \gamma F(V', H') + \frac{1}{2} \|(V', H') - (V, H)\|_F^2. \quad (4.52)$$

En appliquant le critère de Fermat on a alors :

$$\gamma(A^* A(\hat{V} - V_{LR}, \hat{H} - H_{LR})) + (\hat{V} - V, \hat{H} - H) = 0 \quad (4.53)$$

que l'on peut réécrire :

$$(I + \gamma A^* A)(\hat{V}, \hat{H}) = \gamma A^* A(V_{LR}, H_{LR}) + (V, H). \quad (4.54)$$

□

Notons que pour passer de la formulation générale à notre cas d'étude, il suffit de faire passer les hyperparamètres de régularisation λ et α dans l'application linéaire L , explicitement, on peut noter $L = \lambda G_\alpha$. On a alors le résultat suivant qui garantit la convergence des variables itérées.

Théorème 4.4.2 (Théorème 1.b [CP11]). *Si $\sigma\tau\lambda^2\|G_\alpha\|_{op}^2 < 1$, alors*

$$(z_t, u_t) \rightarrow_{t \rightarrow \infty} (z^*, y^*) \quad (4.55)$$

où (z^*, y^*) est un point selle. De plus, si $\delta = \min_z E(z) + P(Lz)$, alors

$$|E(z_t) + P(Lz_t) - \delta| = \mathcal{O}\left(\frac{1}{t}\right). \quad (4.56)$$

Lemme 4.4.3. $\|D_\alpha\|_{op} = \max(1, \alpha)\|D\|_{op}$

Il est en fait possible de proposer une version accélérée de cet algorithme dans le cas où F est fortement convexe en choisissant différents pas σ et τ au cours des itérations. Par exemple, le schéma

$$\begin{cases} z_{t+1} &= \text{prox}_{\sigma F}(z_t - \sigma\lambda D_\alpha^* \tilde{u}_t) \\ u_{t+1} &= \text{prox}_{\tau\|\cdot\|_{2,1}^*}(u_t + \tau\lambda D_\alpha z_{t+1}) \\ \theta_t &= \frac{1}{\sqrt{1+2\gamma\sigma_t}}, \quad \tau_{t+1} = \tau_t\theta_t, \quad \sigma_{t+1} = \sigma_t/\theta_t \\ \tilde{u}_{t+1} &= u_{t+1} + \theta(u_{t+1} - u_t). \end{cases}$$

vérifie le théorème.

Théorème 4.4.4 (Théorème 2 [CP11]). *Si $\sigma_0\tau_0\lambda^2\|D_\alpha\|_{op}^2 < 1$ et $\gamma < \lambda_{\min}(A^* A)$, alors $(z_t, u_t) \rightarrow_{t \rightarrow \infty} (z^*, u^*)$ où (z^*, u^*) est un point selle. De plus,*

$$\|z_t - z^*\| = \mathcal{O}(1/t^2). \quad (4.57)$$

L'intérêt d'avoir une constante de forte convexité⁴ $\min \text{Spec}(A^*A)$ la plus grande possible est que les pas réalisés (τ) dans le dual vont diminuer et être de plus en plus petits (et donc précis). De même, le produit $\sigma_t \tau_t$ ne dépend pas de t , donc les pas (σ) pour l'attache aux données tendent à augmenter. L'interprétation est qu'avec une grande constante de forte convexité, par exemple sur le terme d'attache aux données, alors l'optimisation sur ce terme est facile. Grâce à cela le taux de convergence est quadratique.

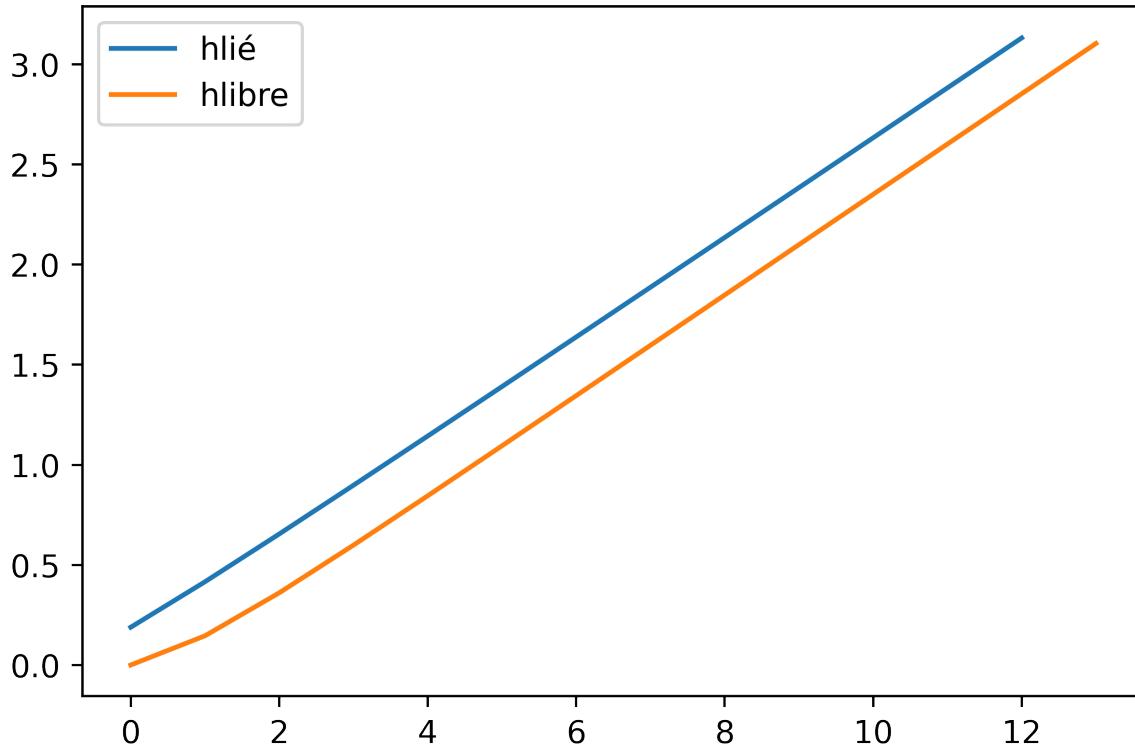


FIGURE 4.3 – Constantes de forte convexité $\min(\text{Spec}A^*A)$ pour les modèles h -lié-min et h -libre-min en fonction du nombre d'échelles disponibles.

4.5 Résultats expérimentaux

Dans cette section, on teste les fonctionnelles proposées sur des champs simulés ainsi que sur des données 'réelles'.

4.5.1 AFBF hétérogène

Les textures considérées dans ces simulations sont obtenues en suivant le modèle des champs browniens fractionnaires anisotropes et simulés à partir de PyAFBF [Ric22]. Afin d'obtenir des textures hétérogènes, on mélange les textures à partir d'un masque, cf fig 4.4. Afin de montrer une vue d'ensemble des résultats, on va tester la capacité à distinguer des

4. Sur la figure 4.3 sont affichées les constantes de forte convexité associées aux modèles h -lié-min et h -libre-min

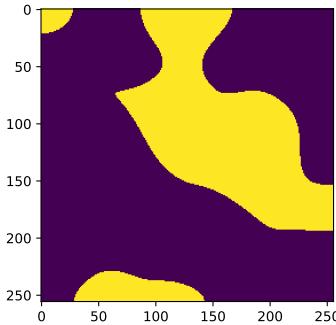


FIGURE 4.4

textures pour lesquelles les caractéristiques (qui permettent la distinction) à extraire sont de plus en plus précises.

Ainsi, on testera la capacité à segmenter deux textures

1. isotropes autosimilaires avec des paramètres de régularité différents,
2. anisotropes autosimilaires,
3. anisotropes non strictement autosimilaires,

Pour tester les capacités de segmentation, les trois tests de segmentations ont été réalisés avec les mêmes paramètres, c'est à dire :

- Des images de taille 256×256
- Une transformée en ondelettes complexes dual tree (cf. section 3.2)
- 7 échelles d'analyse
- Optimisation par un algorithme de Chambolle-Pock avec 1000 itérations
- Aucune calibration précise des hyper-paramètres ($\lambda = 100$ et $\alpha = 1$)

Ces choix ont été faits afin de montrer les capacités de segmentation de la méthode présentée sur des images où assez peu d'échelles d'analyse sont disponibles ($j = 0, \dots, 6$) avec des algorithmes accélérés. De plus, le fait d'obtenir des résultats satisfaisant sans calibration précise des hyperparamètres montre une forme de robustesse des méthodes. Cela motive aussi le fait de développer des procédures efficaces de sélection d'hyperparamètres afin de pouvoir exploiter pleinement les capacités des algorithmes.

Aussi, mentionnons que les deux premières textures homogènes présentées (cf. figures 4.6 et) sont difficiles à distinguer, que ce soit à l'œil ou d'un point de vue statistique. Sur des textures plus faciles à distinguer les résultats auraient été encore plus précis. Aussi, pour ces textures, le choix a été fait d'appliquer le modèle *h-lié-min* qui a été justifié dans la section 4.1.3.

Finalement, les deux dernières textures considérées, bien que faciles à distinguer à l'œil nu, sont difficiles à distinguer avec le modèle *h-lié-min* et la procédure de segmentation échoue. Cela s'explique par le choix des modèles où l'anisotropie est surtout portée par la fonction de Hurst. Dans cette situation, on peut cependant utiliser le modèle *h-libre-min* qui permet alors de segmenter les deux textures.

Aussi, mentionnons que le choix a été fait d'utiliser K -means avec $K = 2$ afin d'obtenir exactement deux régions distinctes. Ce choix n'est pas idéal et pourrait être amélioré. En

particulier, sur plusieurs exemples on peut constater que les zones de discontinuité diminuent la qualité des résultats. Afin d'évaluer les méthodes proposées ici il est donc plus raisonnable de considérer les coefficients régularisés.

En règle générale, le modèle *h-libre-min* est un modèle de segmentation qui fonctionne plus généralement que le modèle *h-lié-min*. Cependant la régularisation se faisant sur deux fois plus de termes, le temps de calcul est doublé. Du point de vue des temps de calcul, sur un ordinateur portable⁵, la segmentation prend environ deux minutes pour *h-lié-min* et quatre minutes pour *h-libre-min*.

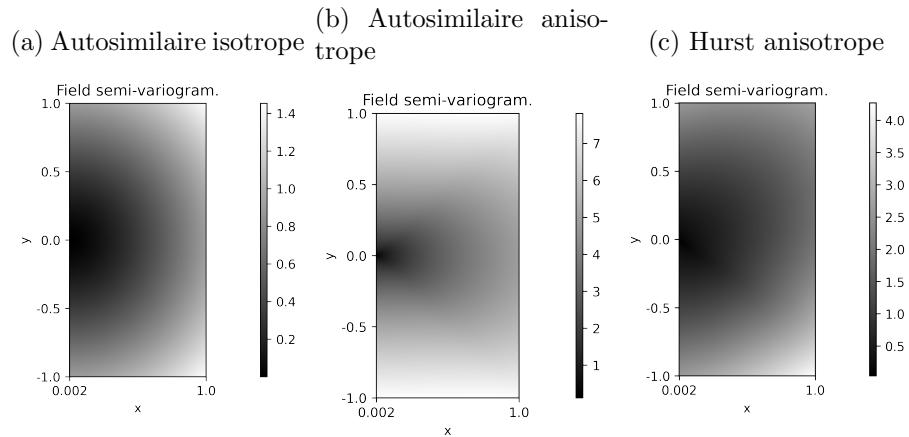


FIGURE 4.5 – Variogrammes d'une des textures homogènes de chacun des modèles considérés

5. Processeur AMD Ryzen 7, 3.2 GHz avec 8Go de RAM

Isotrope h -lié-min

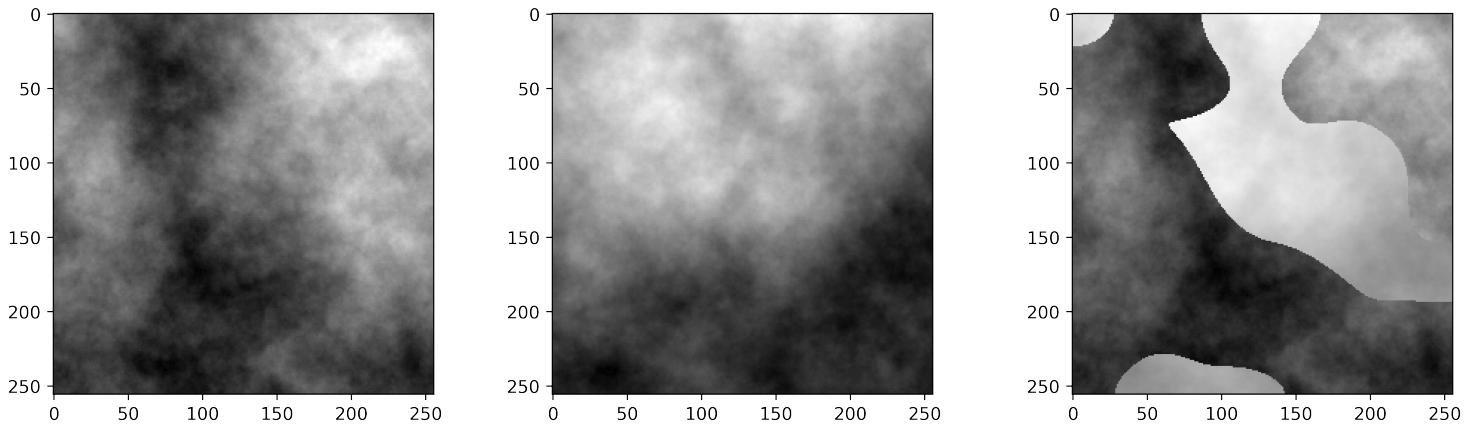


FIGURE 4.6

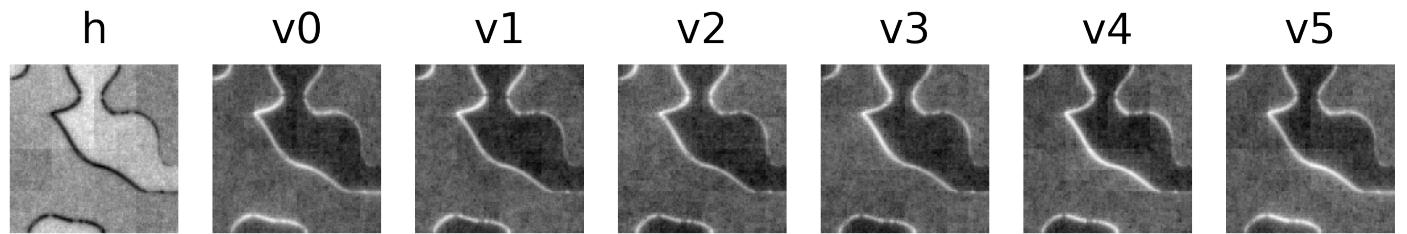


FIGURE 4.7

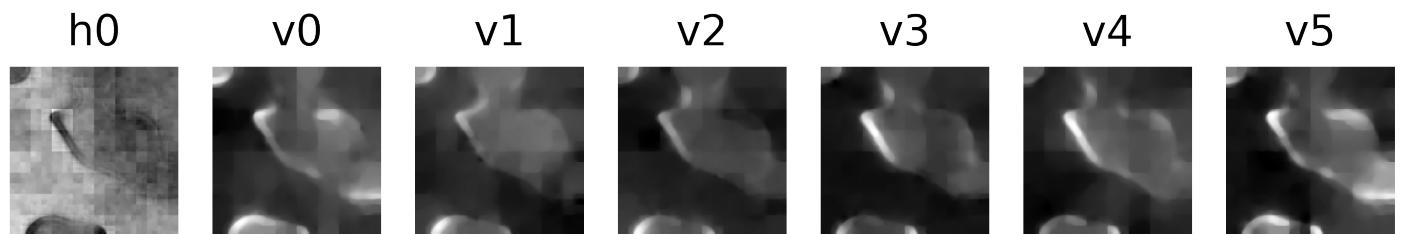


FIGURE 4.8

- Fig. : 4.6 : Textures isotropes de taille 256×256 avec régularité de Hölder égale à 0.7 (Cf. variogramme 4.5a) et 0.65 suivi du mélange des textures par masquage.
- Figs. 4.7 : Coefficients de régression linéaire (H_{LR}, V_{LR}) avant régularisation.
- Figs. 4.8 : Coefficients régularisés par Chambolle-Pock.
- Fig. 4.9 : Image segmentée par 2-means d'après coefficients régularisés, avec pourcentage de pixels mal classifiés.

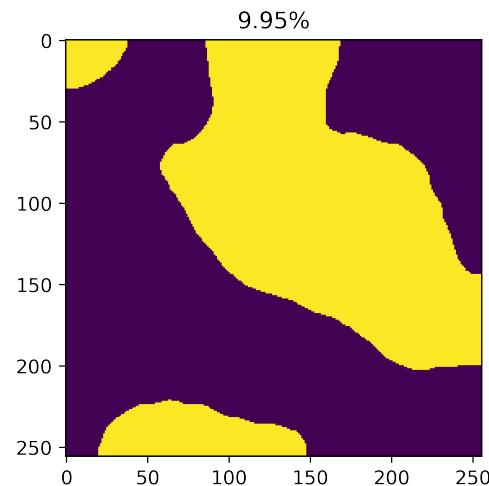


FIGURE 4.9

Autosimilaire anisotrope h -lié-min

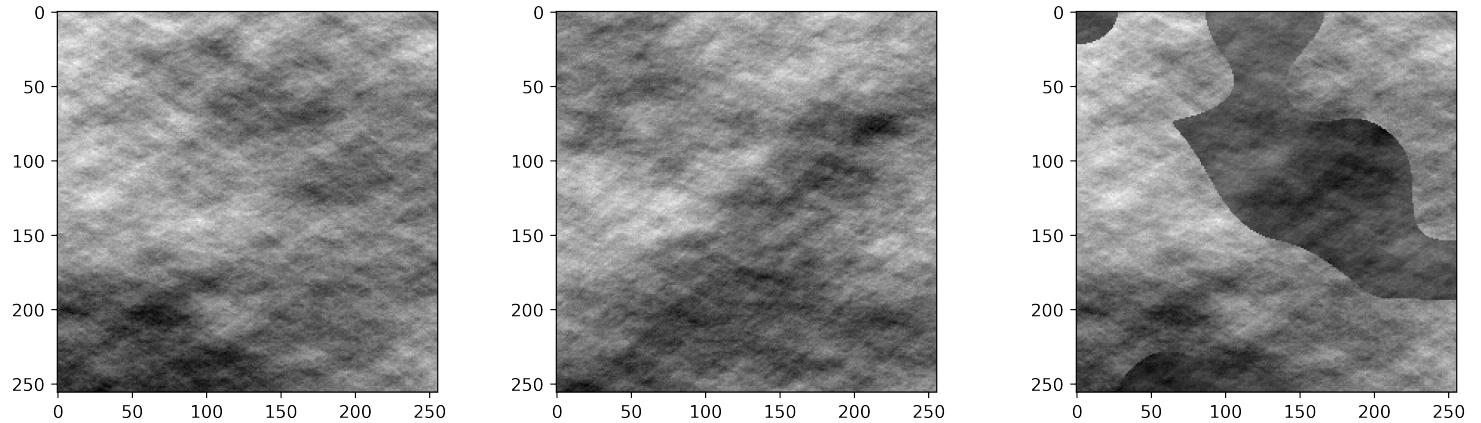


FIGURE 4.10

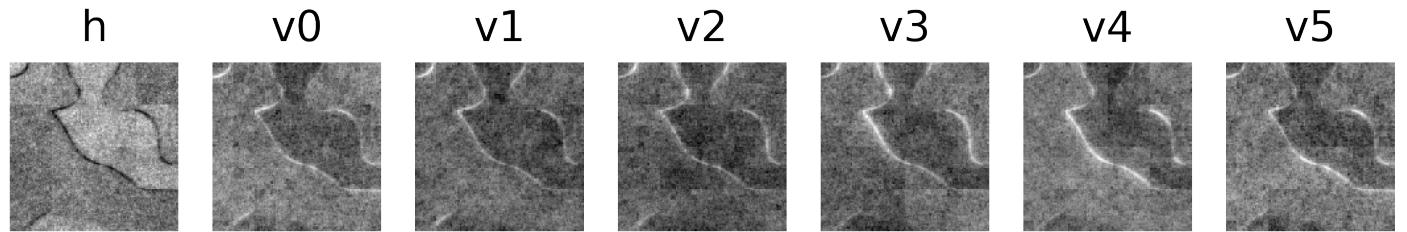


FIGURE 4.11



FIGURE 4.12

- Figs. 4.10 : Textures anisotropes de taille 256×256 avec régularité de Hölder égale à 0.3, ouverture de τ égale à $\pi/2$ prenant les valeurs 1 et 10 (Cf. variogramme 4.5b). Les topothèses τ des deux textures sont déphasées de 10° , suivies du mélange des textures par masquage.
- Fig. 4.11 : Coefficients de régression linéaire (H_{LR}, V_{LR}) avant régularisation.
- Fig. 4.12 : Coefficients régularisés par Chambolle-Pock.
- Fig. 4.13 : Image segmentée avec pourcentage de pixels mal classifiés.

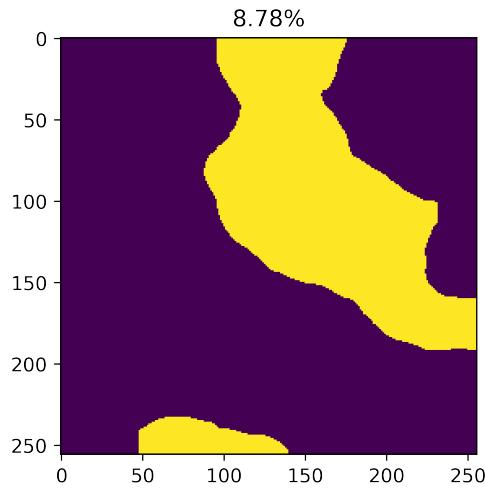


FIGURE 4.13

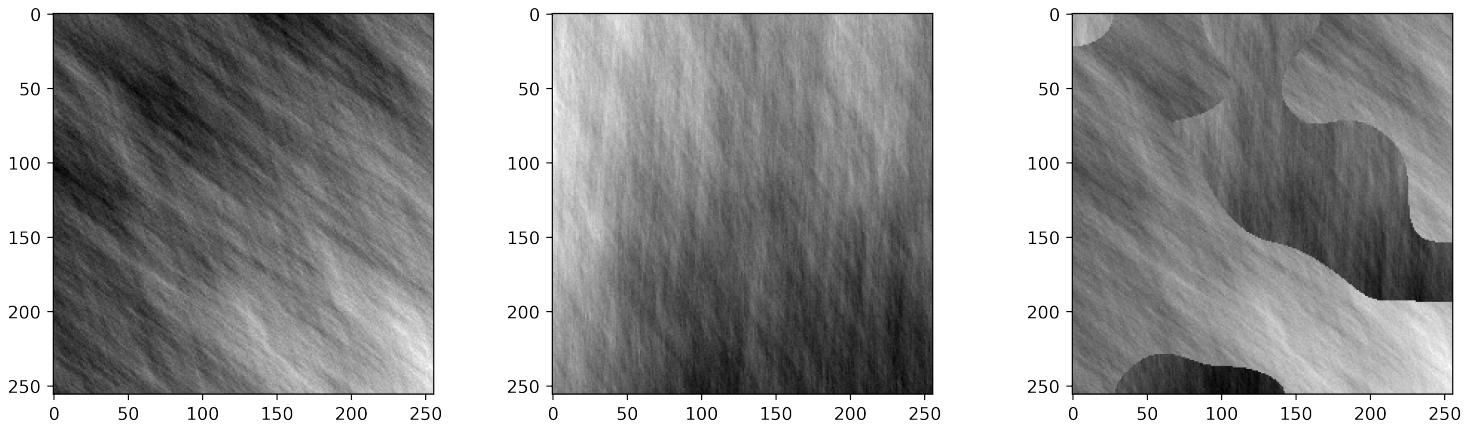


FIGURE 4.14

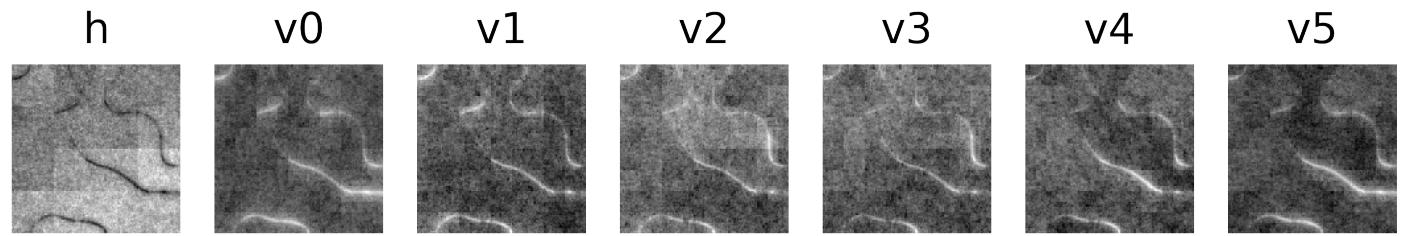


FIGURE 4.15

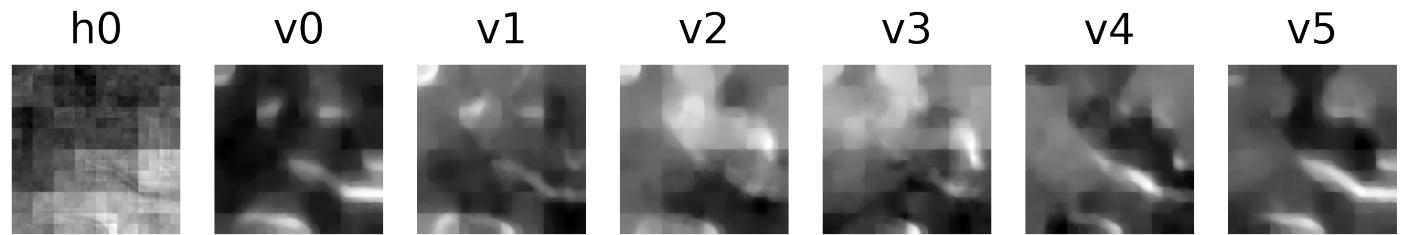


FIGURE 4.16

De haut en bas :

- Figs. 4.14 : Textures anisotropes avec fonction de Hurst (resp. topothèse τ) avec ouverture égale à $\pi/3$ prenant les valeurs 0.8 et 0.2 (resp. 3 et 1) (Cf. variogramme 4.5c). Les fonction de Hurst et de topothèses τ des deux textures sont légèrement déphasées, suivies du mélange des textures par masquage.
- Fig. 4.15 : Coefficients de régression linéaire (H_{LR}, V_{LR}) avant régularisation.
- Fig. 4.16 : Coefficients régularisés par Chambolle-Pock.
- Fig. 4.17 : Image segmentée.

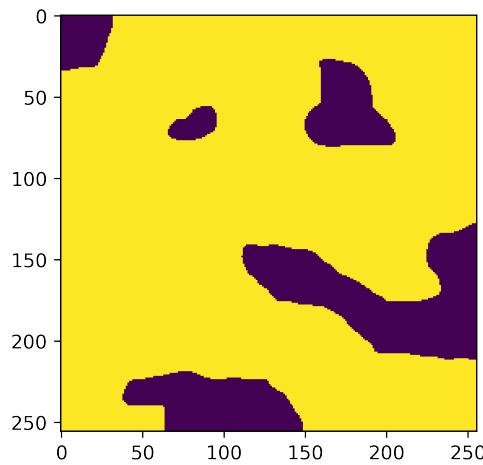


FIGURE 4.17

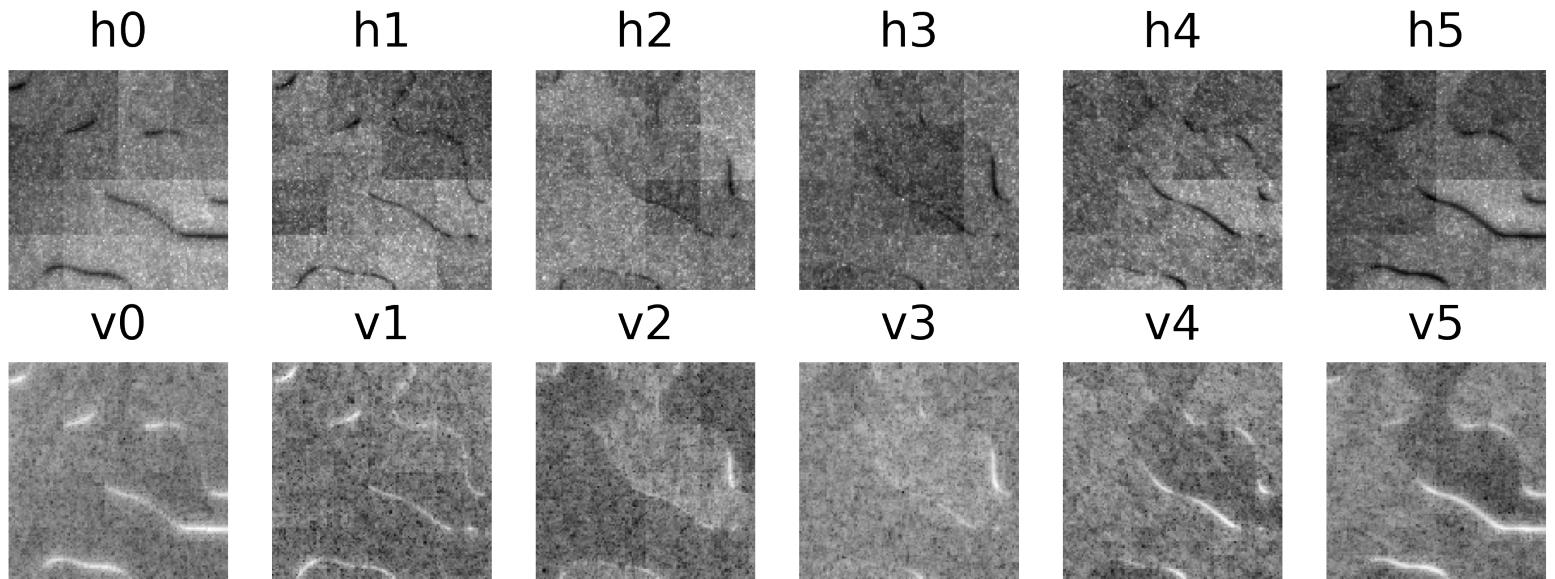


FIGURE 4.18

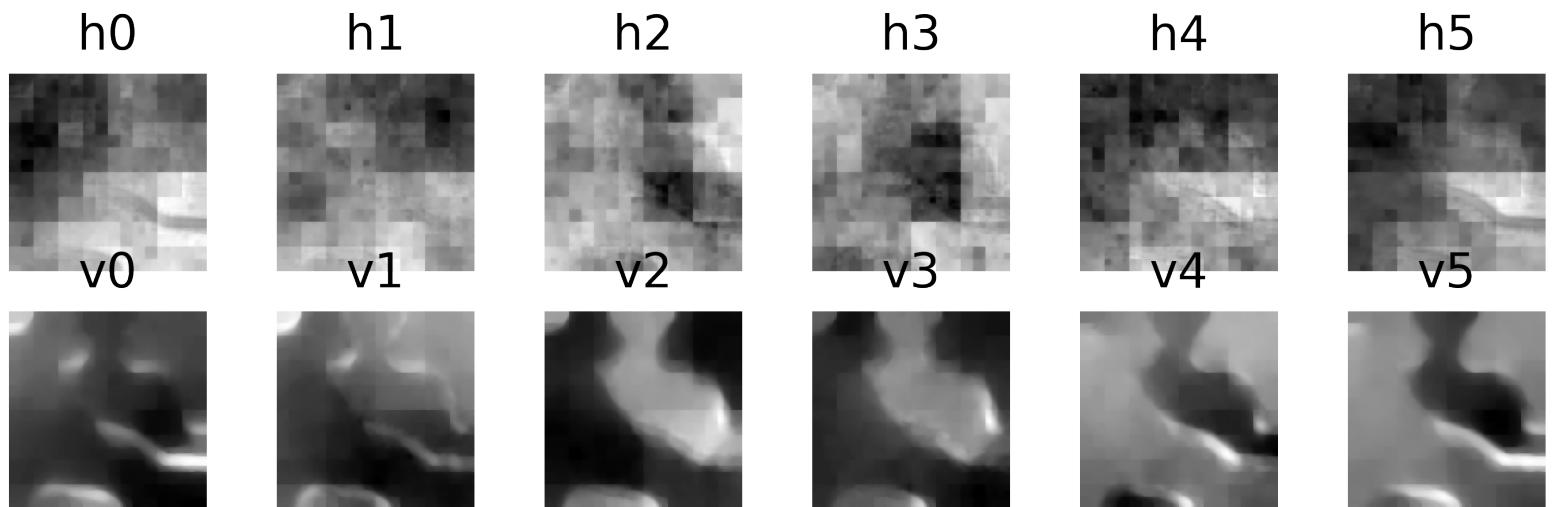


FIGURE 4.19

De haut en bas :

- Fig. 4.18 : Coefficients de régression linéaire (H_{LR}, V_{LR}) du modèle h -libre-min avant régularisation.
- Fig. 4.19 : Coefficients régularisés par Chambolle-Pock.
- Fig. 4.20 : Image segmentée avec pourcentage de pixels mal classifiés.

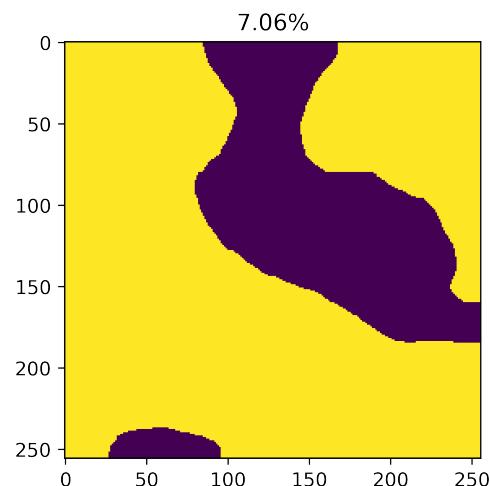


FIGURE 4.20

4.5.2 Brodatz

Le second exemple de textures considérés ici provient des données de Brodatz⁶. Brodatz est un recueil de textures destinés à des artistes et des designers et ne sont pas obtenues selon le modèle AFBF qui a été présenté et qui a été utilisé pour développer les fonctionnelles h -lié-min et h -libre-min. Afin de tester les capacités de segmentation on teste les deux modèles développés sur une mosaïque 4.21a obtenue à l'aide d'un masque 4.21b. Ces textures 'réelles' ont été choisies car elles présentent à la fois des caractéristiques de régularité et de directionnalité.

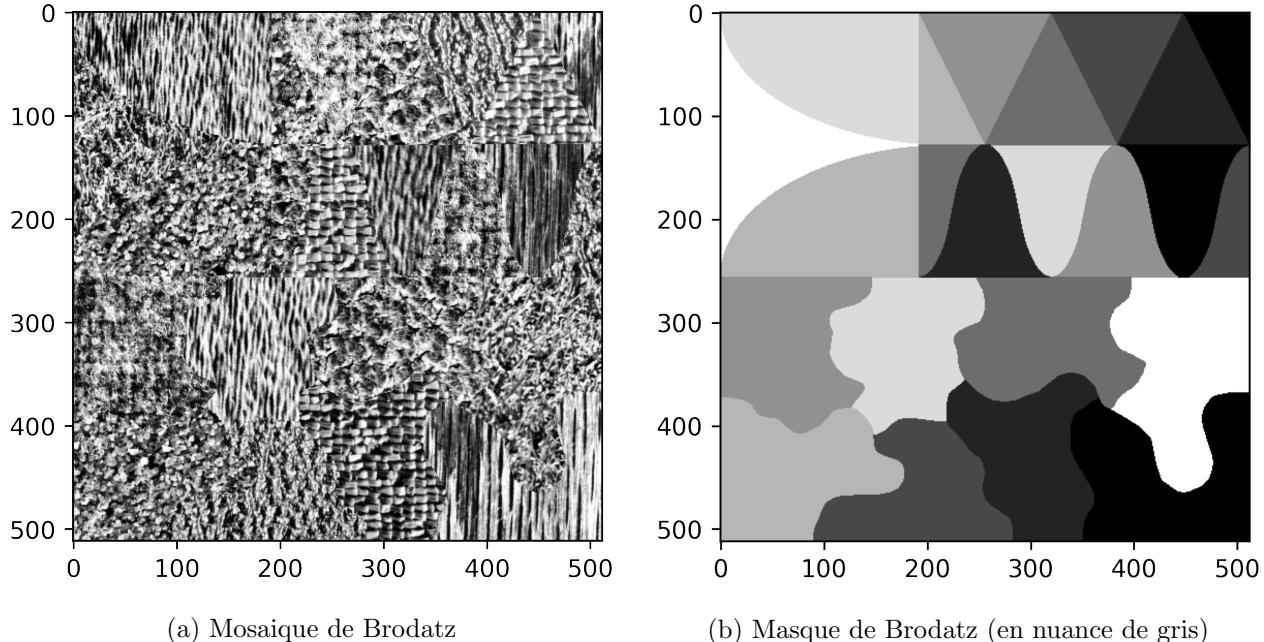


FIGURE 4.21

On réalise la segmentation sur la mosaïque de Brodatz 4.21a avec l'algorithme de Chambolle-Pock avec h -lié-min et h -libre-min, et les mêmes paramètres que précédemment (1000 itérations, $\lambda = 100$ et $\alpha = 1.0$). Les résultats sont présentés aux Figs. 4.22 et 4.23 pour $\lambda = 0$ (avant régularisation). Les résultats régularisés avec $\lambda = 100$ sont présentés aux Figs. 4.24 et 4.25.

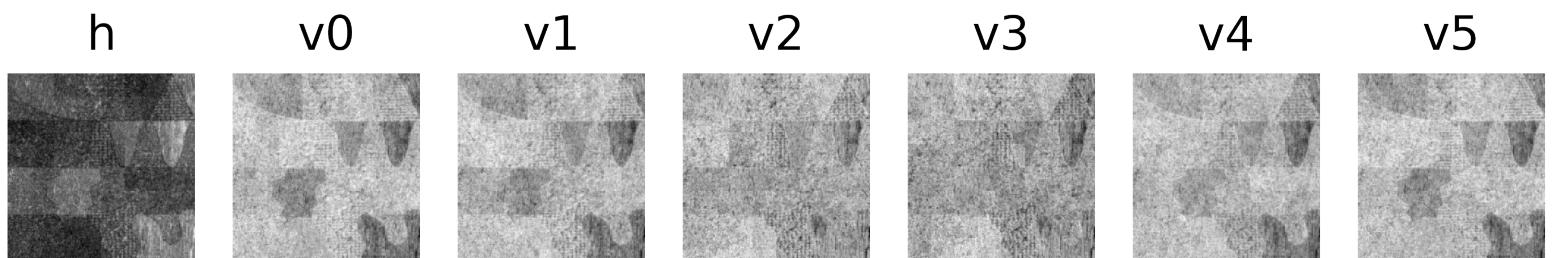


FIGURE 4.22 – Coefficients de régression linéaire non régularisés selon h -lié-min

On remarque sur les images (avant et après) régularisation que les zones constantes par morceaux identifiées sont en assez bonne correspondance avec les zones indiquées sur le masque 4.21b. De plus, nous représentons Fig. 4.26 les segmentations associées par K -means sur les coefficients régularisés avec $K = 8$, conformément au nombre de textures distinctes présentes sur l'image.

6. <https://sipi.usc.edu/database/database.php?volume=textures>

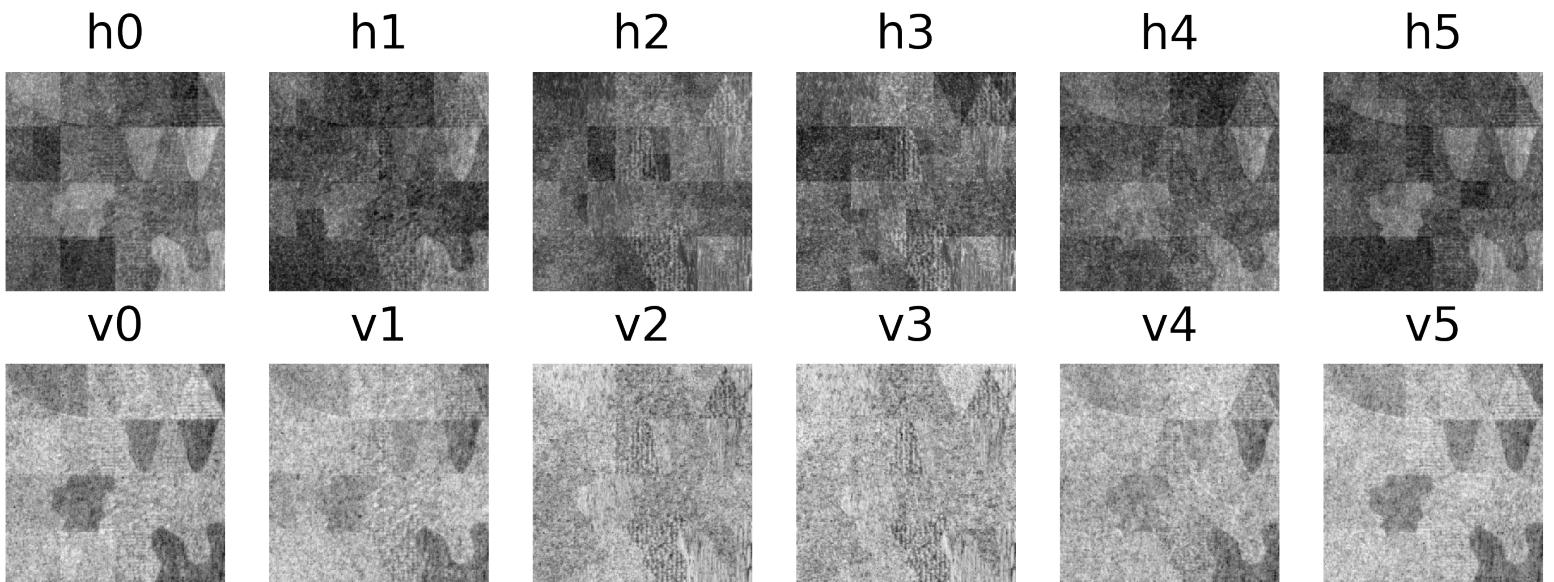


FIGURE 4.23 – Coefficients de régression linéaire non régularisés selon h -libre-min

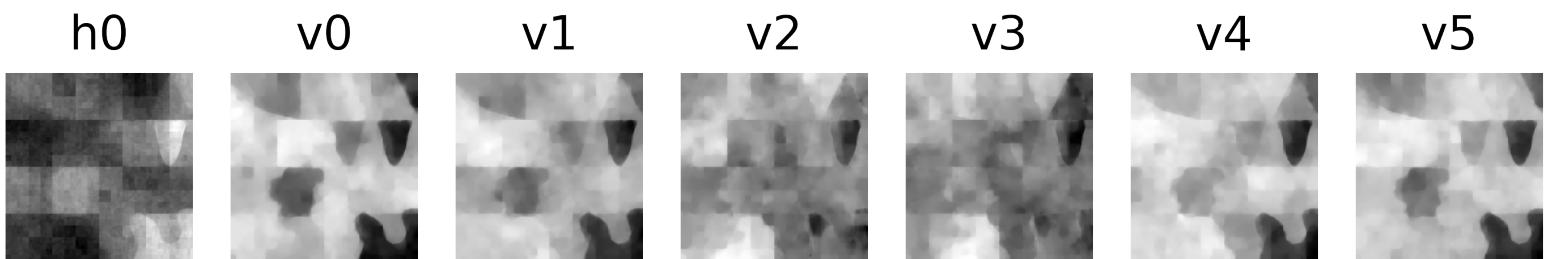


FIGURE 4.24 – Coefficients de régression linéaire après régularisation selon h -lié-min

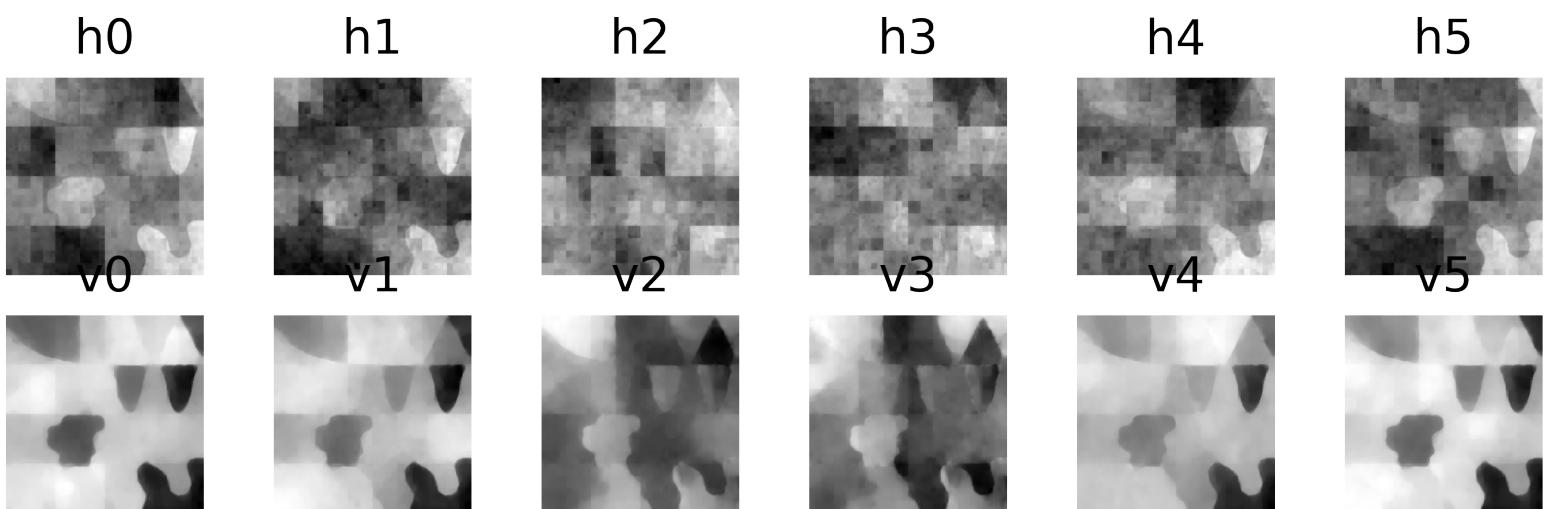
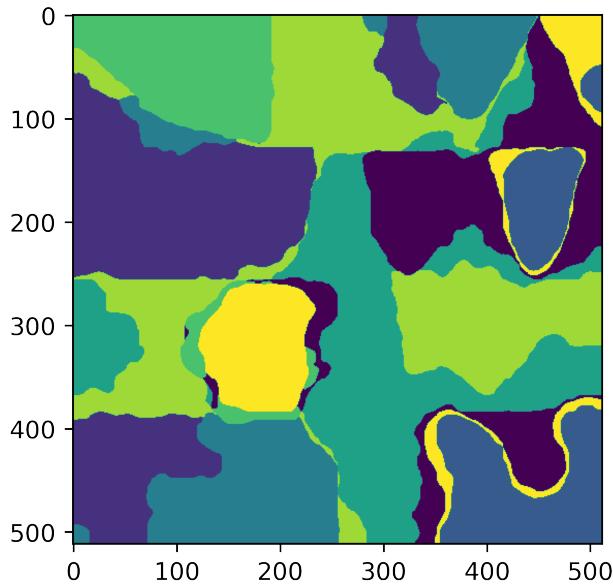
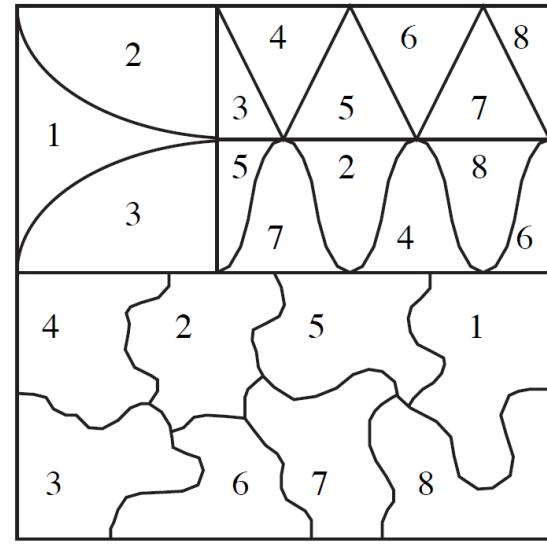


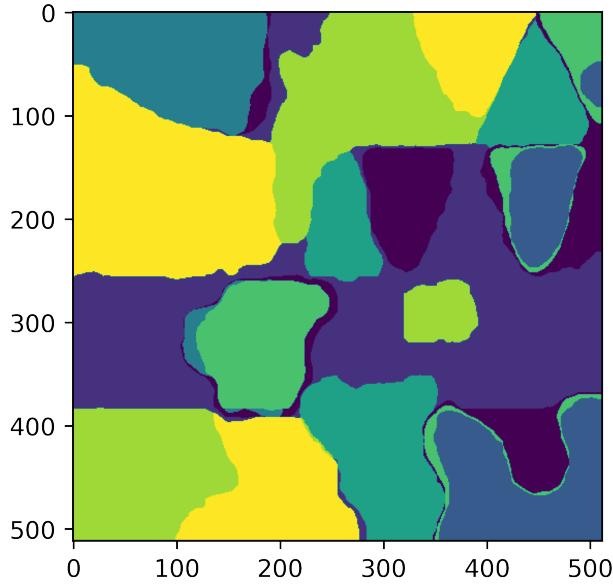
FIGURE 4.25 – Coefficients de régression linéaire après régularisation selon h -libre-min



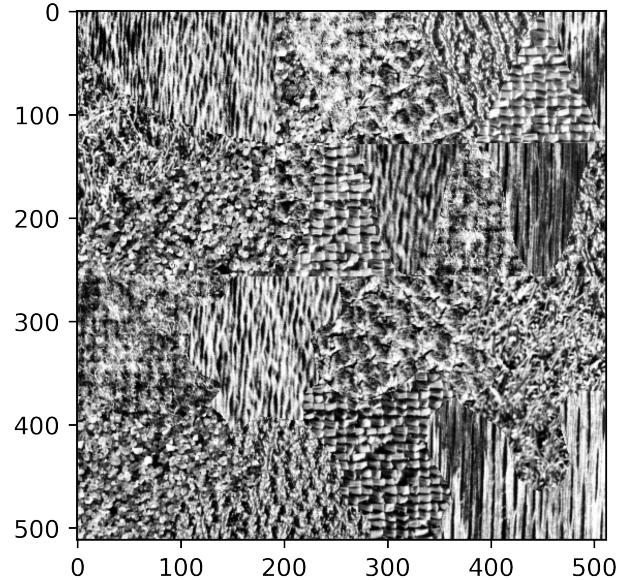
(a) Segmentation K -means d'après h -lié-min avec coefficients régularisés 4.24



(b) Masque de Brodatz



(c) Segmentation K -means d'après h -libre-min avec coefficients régularisés 4.25



(d) Mosaique de Brodatz

FIGURE 4.26

Chapitre 5

Conclusion et perspectives

En conclusion, nous avons considéré le problème de segmentation de textures irrégulières et on s'est en particulier concentré sur l'influence de paramètres directionnels.

Du point de vue de la *modélisation* nous avons considéré le modèle des champs browniens fractionnaires anisotropes. Pour ces modèles nous avons vu un ensemble de résultats dus à Aline Bonami et Anne Estrade permettant de caractériser les propriétés des réalisations à partir de propriétés de la mesure spectrale associée à la représentation harmonisable des modèles.

Dans des travaux futurs nous pourrons considérer les extensions suivantes :

1. Des modèles plus généraux de textures pourront être considérés. La notion plus générale d'autosimilarité aurait pu être approfondie, soit par des notions fractales [Abr+15] [Xia13], soit en remplaçant le paramètre réel d'autosimilarité par un paramètre matriciel [BDP14].
2. Une autre possibilité de développement peut se faire par l'étude de modèles hétérogènes, c'est à dire, pour lesquels les caractéristiques dépendent continûment de la position. De plus, cela permettrait la synthèse de textures n'ayant pas les défauts de continuités qui apparaissent en utilisant des masques présentés ici[Pol+14].

Nous avons ensuite considéré une méthode d'analyse multi-échelle bénéficiant de nombreuses bonnes propriétés, qui proviennent à la fois de la théorie des ondelettes et de la transformée de Hilbert avec la dual-tree complex wavelet transform de Nick Kingsbury. En combinant cette analyse à un formalisme multi-échelle nous avons pu vérifier la validité de cette approche sur des textures homogènes.

Dans des travaux futurs nous envisagerons de poursuivre les travaux initiés des façons suivantes :

1. Du point de vue de l'analyse, la première chose à approfondir est le lien entre les propriétés spectrales et l'analyse spatiale qui est réalisée. Pour cela un traitement détaillé de la quantité stochastique $\langle X^f, \psi_{j,k} \rangle$, notamment par rapport aux propriétés spectrales de ψ serait d'intérêt à la fois pour la description des modèles et aussi pour le choix des filtres ψ à utiliser. Aller au delà de l'analyse en ondelettes par une transformée en scattering par exemple pourrait également être pertinent[BM13].
2. Une autre quantité importante à développer est le tenseur de structure $c_{j,k}c_{j,k}^*$, où on

note $c_{j,k} = (c_{j,k}^{(d)})_{d=1}^D$ les coefficients à travers les directions. L'analyse développée ici s'intéresse seulement aux termes diagonaux du tenseur de structure. Cependant, les termes hors diagonale permettent de prendre en compte des informations sur la covariance des coefficients, qui peut notamment être contenue dans les valeurs/vecteurs propres [Ore+22][Abr+21]. Aussi, en traitement des images, considérer ce type de quantité permet de prendre en compte des informations telles que les couleurs de l'image et d'améliorer les résultats [Chi+14].

Finalement, nous avons introduit plusieurs fonctionnelles et algorithmes permettant la prise en compte des hypothèses de modélisation et d'homogénéité locale exprimables avec des opérateurs proximaux. Nous avons ainsi pu développer des algorithmes adaptés à notre modèle pour lesquels nous avons présenté des versions accélérées, par exemple grâce à la forte convexité de l'attache aux données.

Pour l'ensemble des méthodes proposées nous avons pu vérifier empiriquement leur validité et une bibliothèque Python a été développée pour permettre une implémentation facile des méthodes. On a alors pu vérifier que les algorithmes proposés permettent la segmentation de textures anisotropes issues d'un champ brownien fractionnaire anisotrope. On a de plus vu l'intérêt des deux modèles (h -libre-min et h -lié-min) en fonction des textures considérées. De plus, quelques tests ont été menés sur des textures hétérogènes ne vérifiant pas le modèle considéré et on a aussi pu constater l'efficacité des méthodes proposées, ce qui suggère une forme de robustesse de la méthodologie proposée.

Dans des travaux futurs, on considérera les différentes perspectives suivantes :

1. La première façon d'améliorer les résultats présentés ici concerne la sélection des hyperparamètres. Le choix de ces hyperparamètres est capital et les temps de fonctionnement des algorithmes proposés rendent même la sélection à posteriori très longue. Des procédures efficaces par descente de gradient à l'aide d'un estimateur du risque non-biaisé de Stein a déjà été développé dans un cas similaire [Pas20][Pas+21] et devrait pouvoir être adapté à la situation considérée.
2. Une autre façon d'améliorer ce qui a été fait ici peut consister en la prise en compte plus tôt dans le traitement du fait que l'objectif est d'avoir une segmentation en un nombre discret de labels. Ici, on a fait cela de façon *ad-hoc* en appliquant un K -means, en sachant à priori la valeur de K . Des méthodes inspirées du compressed-sensing pour la reconstruction parcimonieuse de mixtures de distributions pourraient également fournir de bons résultats[Gri+20].
3. Une autre piste avec un fort potentiel qui sera fortement étudiée pendant la thèse est celle basée sur les algorithmes déroulés. Pendant ces dernières années, il est devenu fréquent d'approcher les problèmes en traitement d'images avec des réseaux de neurones ou des stratégies analogues, motivés essentiellement par un succès empirique. Ces réseaux, avec peu de couches, sont quasiment entièrement déduits des données (*data driven*).

Au contraire, la stratégie présentée ici est justifiée par des théories mathématiquement solides (processus stochastique, analyse fonctionnelle, analyse convexe,...) mais nécessite un grand nombre d'itérations (*theory driven*).

L'étude des algorithmes déroulés peut prendre son point de départ de l'observation que les algorithmes obtenus en général à partir de la théorie des opérateurs proximaux ont une structure très similaire à des réseaux de neurones classiques. Une cascade

d'opérations non-linéaires simples et d'applications linéaires. Nous serons ainsi amenés à considérer les algorithmes déroulés en tant que passerelle d'une approche à l'autre. L'objectif à terme, étant de voir si, et par quel moyen, il est possible de faire passer des propriétés d'une approche à l'autre[LPF22][Che+18].

Bibliographie

- [Abr+15] Patrice ABRY et al. “The hyperbolic wavelet transform : an efficient tool for multi-fractal analysis of anisotropic fields”. In : *Revista matemática iberoamericana* 31.1 (2015), p. 313-348.
- [Abr+21] Patrice ABRY et al. “Wavelet eigenvalue regression in high dimensions”. In : *arXiv preprint arXiv :2108.03770* (2021).
- [BC+11] Heinz H BAUSCHKE, Patrick L COMBETTES et al. *Convex analysis and monotone operator theory in Hilbert spaces*. T. 408. Springer, 2011.
- [BDP14] Changryong BAEK, Gustavo DIDIER et Vladas PIPIRAS. “On integral representations of operator fractional Brownian fields”. In : *Statistics & Probability Letters* 92 (2014), p. 190-198.
- [BE03] Aline BONAMI et Anne ESTRADE. “Anisotropic analysis of some Gaussian models”. In : *Journal of Fourier analysis and applications* 9.3 (2003), p. 215-236.
- [Bie05] Hermine BIERMÉ. “Champs aléatoires : autosimilarité, anisotropie et étude directionnelle”. Theses. Université d’Orléans, 2005. URL : <https://tel.archives-ouvertes.fr/tel-00010227>.
- [Bie19] Hermine BIERMÉ. “Introduction to Random Fields and Scale Invariance”. In : *Stochastic Geometry : Modern Research Frontiers*. Sous la dir. de David COUPIER. Springer International Publishing, 2019, p. 129-180.
- [BM13] Joan BRUNA et Stéphane MALLAT. “Invariant scattering convolution networks”. In : *IEEE transactions on pattern analysis and machine intelligence* 35.8 (2013), p. 1872-1886.
- [BMR15] Hermine BIERMÉ, Lionel MOISAN et Frédéric RICHARD. “A turning-band method for the simulation of anisotropic fractional Brownian fields”. In : *Journal of Computational and Graphical Statistics* 24.3 (2015), p. 885-904.
- [CD15] Antonin CHAMBOLLE et Charles H DOSSAL. “On the convergence of the iterates of “FISTA””. In : *Journal of Optimization Theory and Applications* Volume 166.Issue 3 (2015), p. 25. URL : <https://hal.inria.fr/hal-01060130>.
- [Che+18] Xiaohan CHEN et al. “Theoretical linear convergence of unfolded ISTA and its practical weights and thresholds”. In : *Advances in Neural Information Processing Systems* 31 (2018).
- [Chi+14] Giovanni CHIERCHIA et al. “A nonlocal structure tensor-based approach for multi-component image recovery problems”. In : *IEEE transactions on Image Processing* 23.12 (2014), p. 5531-5544.

- [Chi+17] G CHIERCHIA et al. *The proximity operator repository (user's guide)*. 2017.
- [CP11] Antonin CHAMBOLLE et Thomas POCK. “A first-order primal-dual algorithm for convex problems with applications to imaging”. In : *Journal of mathematical imaging and vision* 40.1 (2011), p. 120-145.
- [Dav21] Leo DAVY. *Mémoire de M1, Reconstruction parcimonieuse des signaux, Frames, décomposition atomique et introduction au compressed sensing, dirigé par Jean-François Crouzet*. 2021. URL : <https://github.com/DavyL/MemoireM1/raw/main/main.pdf>.
- [Gri+20] Rémi GRIBONVAL et al. “Sketching datasets for large-scale learning (long version)”. In : *arXiv preprint arXiv:2008.01839* (2020).
- [Jaf+16] Stephane JAFFARD et al. “p-exponent and p-leaders, Part I : Negative pointwise regularity”. In : *Physica A : Statistical Mechanics and its Applications* 448 (2016), p. 300-318.
- [Jan+21] Ryoungwoo JANG et al. “Prediction of osteoporosis from simple hip radiography using deep learning algorithm”. In : *Scientific Reports* 11 (oct. 2021), p. 19997. DOI : 10.1038/s41598-021-99549-6.
- [Kin06] Nick KINGSBURY. “Rotation-invariant local feature matching with complex wavelets”. In : *2006 14th European Signal Processing Conference*. IEEE. 2006, p. 1-5.
- [LG13] Jean-François LE GALL. *Mouvement brownien, martingales et calcul stochastique*. Springer, 2013.
- [LPF22] Hoangtrieuvy LE, Nelly PUSTELNIK et Marion FOARE. “The faster proximal algorithm, the better unfolded deep learning architecture ? The study case of image denoising”. In : (2022).
- [Mal89] Stephane G MALLAT. “A theory for multiresolution signal decomposition : the wavelet representation”. In : *IEEE transactions on pattern analysis and machine intelligence* 11.7 (1989), p. 674-693.
- [Man75] Benoît B MANDELBROT. *Les objets fractals : forme, hasard et dimension*. T. 17. Flammarion Paris, 1975.
- [Ore+22] Oliver OREJOLA et al. “Hurst multimodality detection based on large wavelet random matrices”. In : *European Signal Processing Conference (EUSIPCO)*. Belgrade, Serbia, 2022. URL : <https://hal.archives-ouvertes.fr/hal-03735495>.
- [Pas+18] Barbara PASCAL et al. “Joint Estimation of Local Variance and Local Regularity for Texture Segmentation. Application to Multiphase Flow Characterization”. In : *2018 IEEE International Conference on Image Processing, ICIP 2018, Athens, Greece, October 7-10, 2018*. IEEE, 2018, p. 2092-2096. URL : <https://doi.org/10.1109/ICIP.2018.8451380>.
- [Pas20] Barbara PASCAL. “Estimation régularisée d'attributs fractals par minimisation convexe pour la segmentation de textures : formulations variationnelles conjointes, algorithmes proximaux rapides et sélection non supervisée des paramètres de régularisation ; Applications à l'étude du frottement solide et de la microfluidique des écoulements multiphasiques”. Thèse de doct. Université de Lyon, 2020.

- [Pas+21] Barbara PASCAL et al. “Automated data-driven selection of the hyperparameters for total-variation-based texture segmentation”. In : *Journal of Mathematical Imaging and Vision* 63.7 (2021), p. 923-952.
- [PC17] Nelly PUSTELNIK et Laurent CONDAT. “Proximity operator of a sum of functions ; application to depth map estimation”. In : *IEEE Signal Processing Letters* 24.12 (2017), p. 1827-1831.
- [Pol+14] Kévin POLISANO et al. “Texture Modeling by Gaussian fields with prescribed local orientation”. In : *2014 IEEE International Conference on Image Processing (ICIP)*. IEEE. 2014, p. 6091-6095.
- [Pol17] Kevin POLISANO. “Modélisation de textures anisotropes par la transformée en ondelettes monogéniques”. Thèse de doct. Université Grenoble Alpes, 2017. URL : <https://tel.archives-ouvertes.fr/tel-01728616>.
- [PPA21] Barbara PASCAL, Nelly PUSTELNIK et Patrice ABRY. “Strongly convex optimization for joint fractal feature estimation and texture segmentation”. In : *Applied and Computational Harmonic Analysis* 54 (2021), p. 303-322.
- [RAL07] Rangaraj M. RANGAYYAN, Fábio J. AYRES et J.E. LEO DESAUTELS. “A review of computer-aided diagnosis of breast cancer : Toward the detection of subtle signs”. In : *Journal of the Franklin Institute* 344.3 (2007). Special Issue : Medical Applications of Signal Processing, Part I, p. 312-348. ISSN : 0016-0032. DOI : <https://doi.org/10.1016/j.jfranklin.2006.09.003>. URL : <https://www.sciencedirect.com/science/article/pii/S001600320600127X>.
- [Ric15] F.J.P. RICHARD. “Analysis of Anisotropic Brownian Textures and Application to Lesion Detection in Mammograms”. In : *Procedia Environmental Sciences* 27 (2015). Spatial Statistics conference 2015, p. 16-20. ISSN : 1878-0296. DOI : <https://doi.org/10.1016/j.proenv.2015.07.100>. URL : <https://www.sciencedirect.com/science/article/pii/S1878029615003126>.
- [Ric16] Frédéric JP RICHARD. “Tests of isotropy for rough textures of trended images”. In : *Statistica Sinica* (2016), p. 1279-1304.
- [Ric18] Frédéric JP RICHARD. “Anisotropy of Hölder Gaussian random fields : characterization, estimation, and application to image textures”. In : *Statistics and Computing* 28.6 (2018), p. 1155-1168.
- [Ric22] Frédéric JP RICHARD. “PyAFBF : a Python library for sampling image textures from the anisotropic fractional Brownian field.” In : *Journal of Open Source Software* 7.75 (2022), p. 3821.
- [SBK05] Ivan W SELESNICK, Richard G BARANIUK et Nick C KINGSBURY. “The dual-tree complex wavelet transform”. In : *IEEE signal processing magazine* 22.6 (2005), p. 123-151.
- [Xia13] Yimin XIAO. “Recent developments on fractal properties of Gaussian random fields”. In : *Further Developments in Fractals and Related Fields* (2013), p. 255-288.