



Ondelette : Tatouage et Stéga-analyse

Philippe Carré

philippe.carre@univ-poitiers.fr

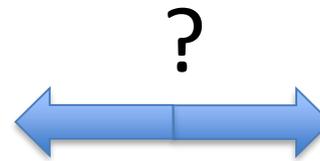
XLIM, UMR 7252, Université de Poitiers





Proposition

Ondelettes



Sécurité



Décompositions atomiques discrètes reconstructibles (ondelettes)

Objet d'étude

$$\begin{aligned} A : [i, j] &\mapsto (A_1, A_2, A_3) \\ \mathbb{Z} \times \mathbb{Z} &\rightarrow (\mathbb{R}, \mathbb{R}, \mathbb{R}) \end{aligned}$$

Outil

$$s = \sum_{k_1 \in \mathbb{Z}} \sum_{k_2 \in \mathbb{Z}} c_{0,k_1,k_2} \zeta_{0,k_1,k_2}$$

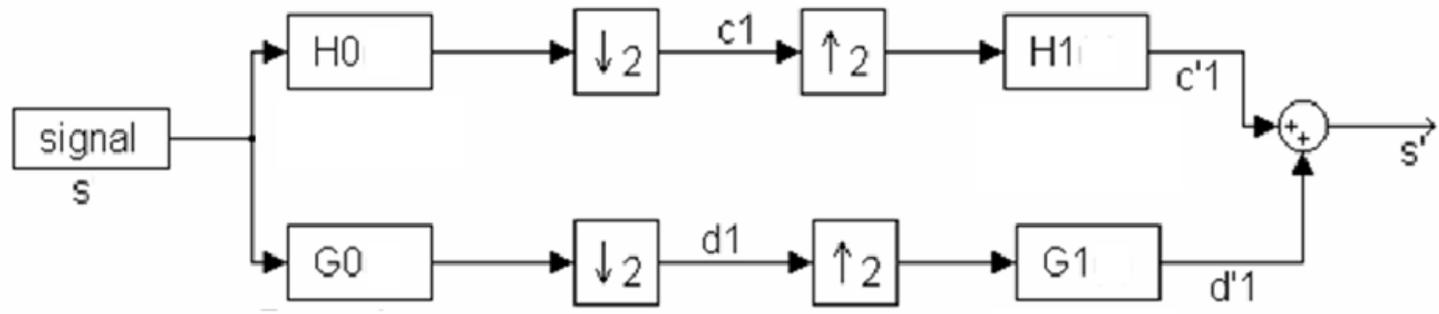
Domaines

Manipulation des coefficients

↓
Sécurité

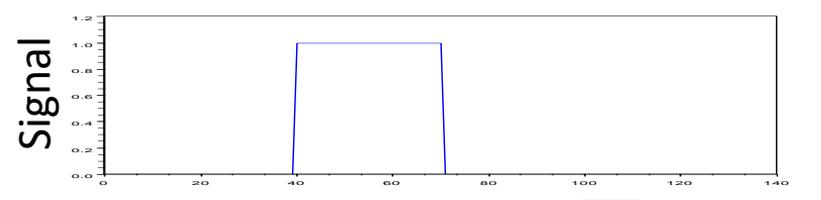
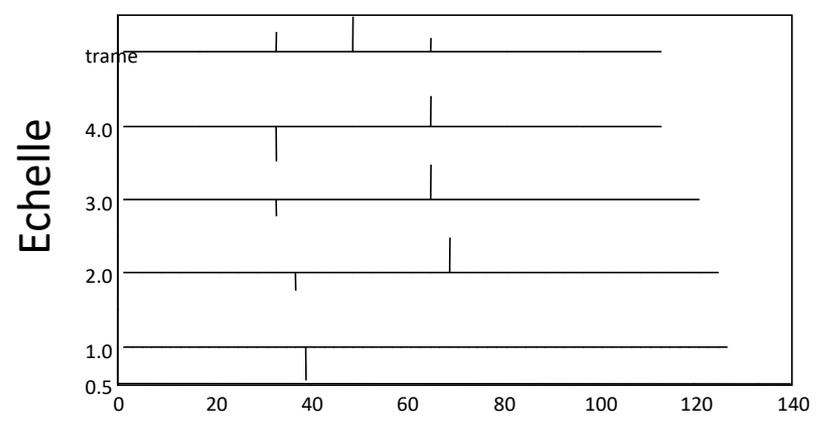


Analyse 1-D scalaire



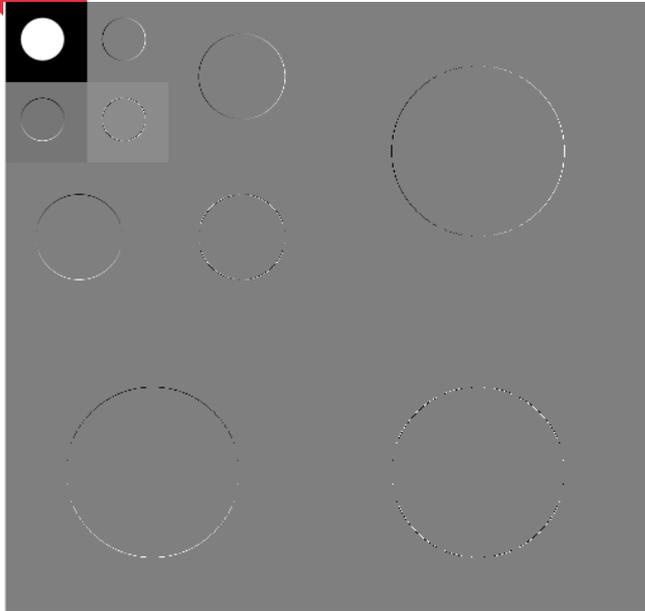
$$\log_2(N) \rightarrow \mathcal{O}(\log_2(N)) < \mathcal{O}(N)$$

- Rapide
- Propriété d'orthogonalité
- Parfaitement restructurable



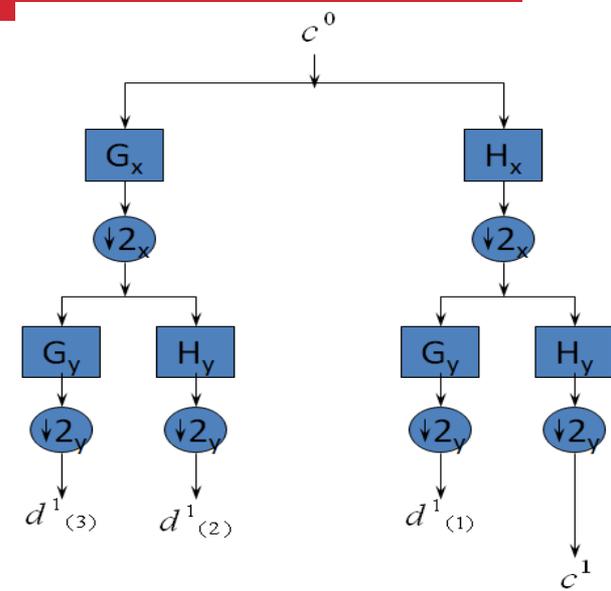
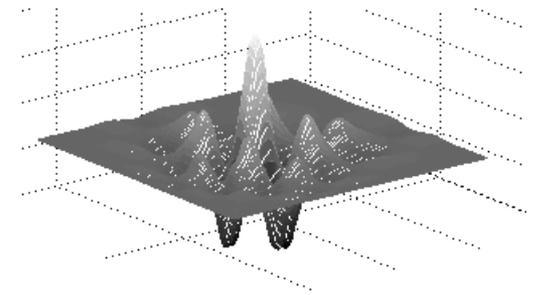
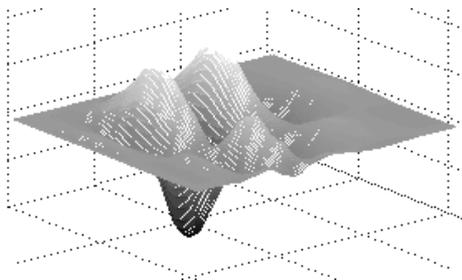
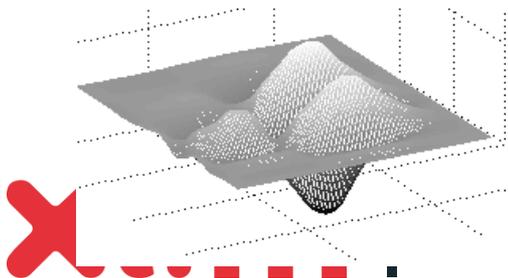
Vers le 2-D scalaire

Discontinuités 2-D



Par échelle

$\mathcal{O}(2^{-l} * lg)$ avec lg la longueur du contour



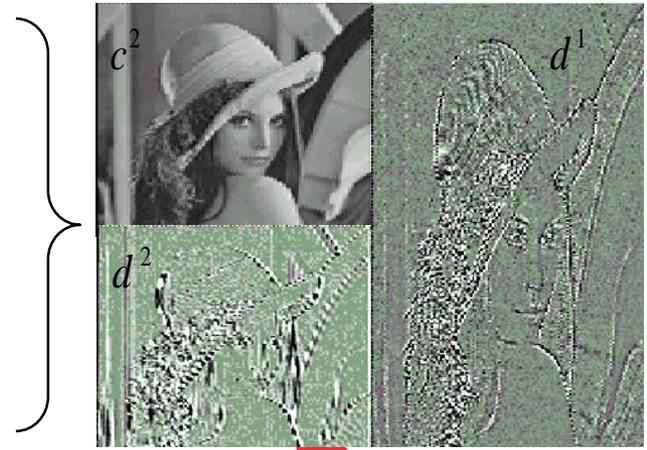
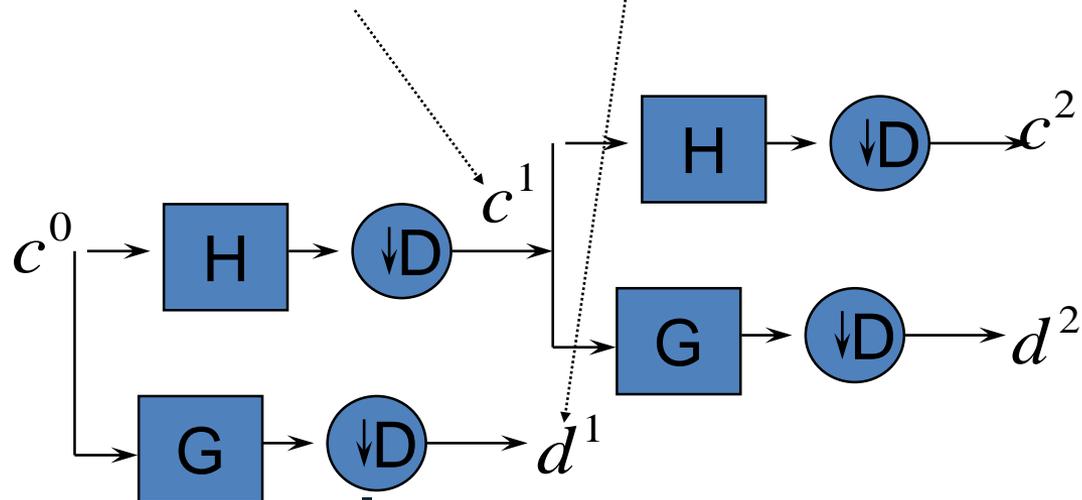
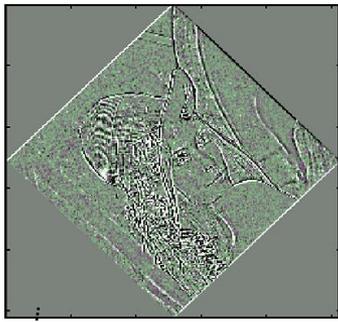
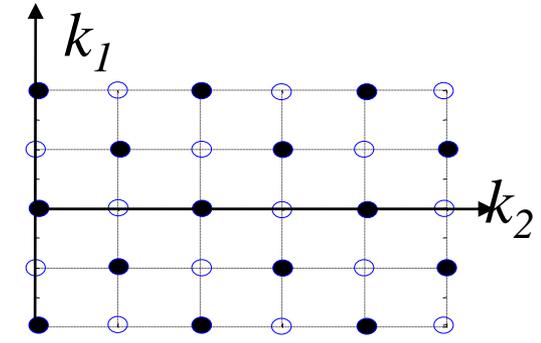
Ondelettes meilleures que Fourier mais

- Détectent les singularités pas la structure contour
- Directions privilégiées
- Trois plans d'ondelettes par échelle

Version non séparable

Dilatation en quinconce

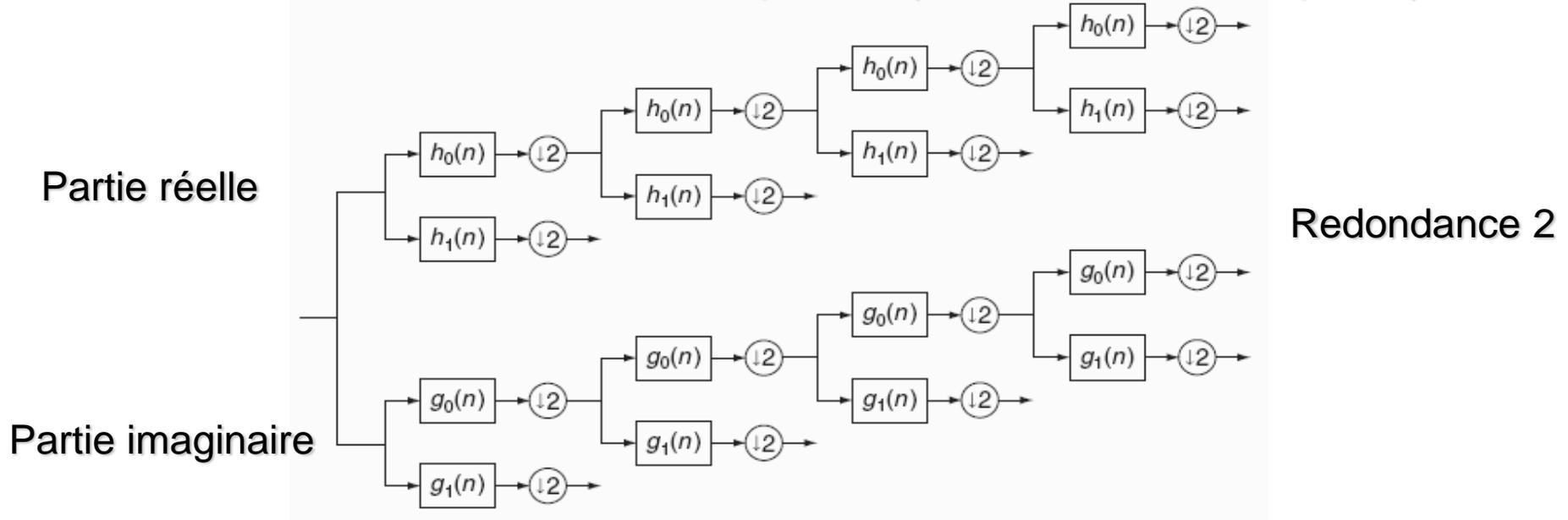
$$D = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$



Une transformée complexe

Solution de Kingsbury et al. : calcul par arbre dual

2 bancs de filtres à reconstruction parfaite (et conditions «classiques »)



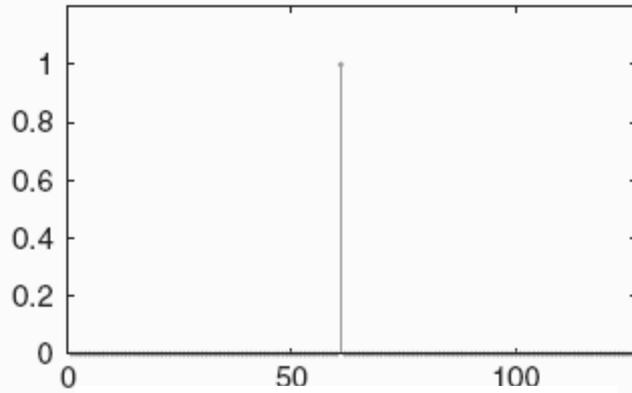
Contrainte : les filtres doivent être définis afin qu'ils vérifient

$$\psi(t) = \psi_h(t) + j\psi_g(t)$$

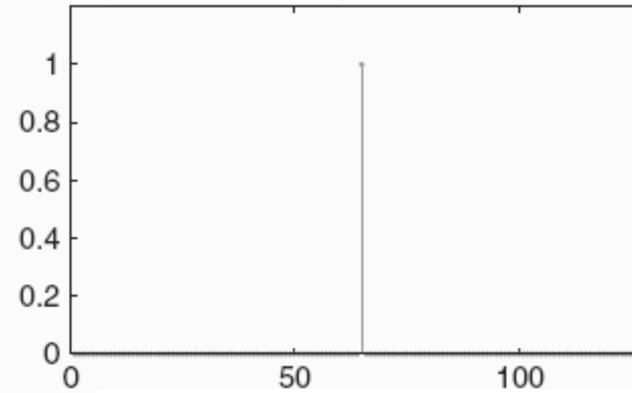




Signal de test $x(n) = \delta(n-60)$

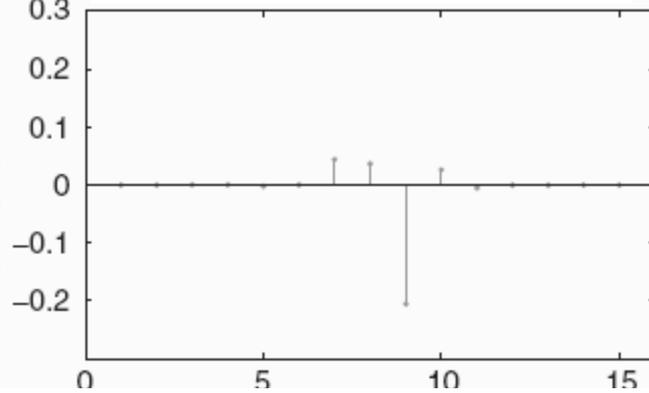


Signal de test $x(n) = \delta(n-64)$

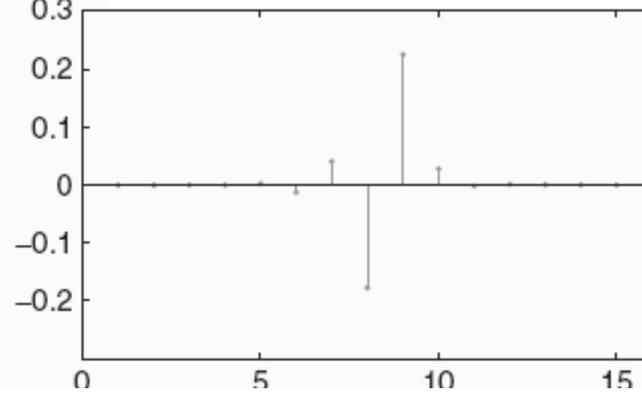


DWT réelle

Première Echelle

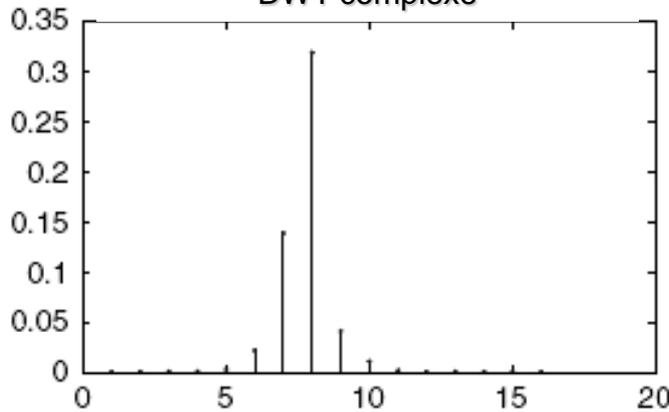


DWT réelle

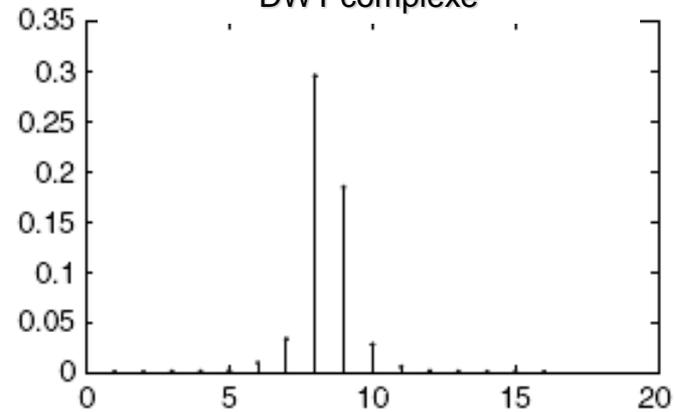


DWT complexe

Première Echelle



DWT complexe

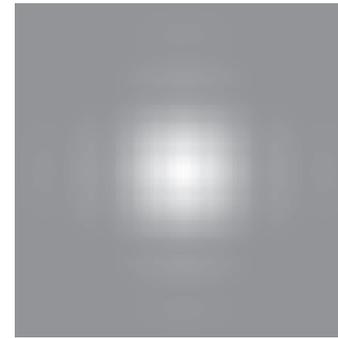
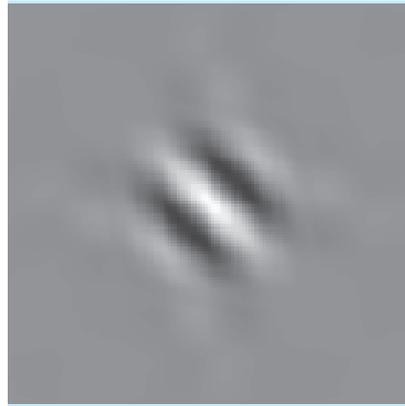
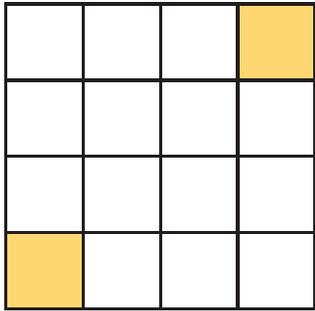


DT 2D : analyse orientée

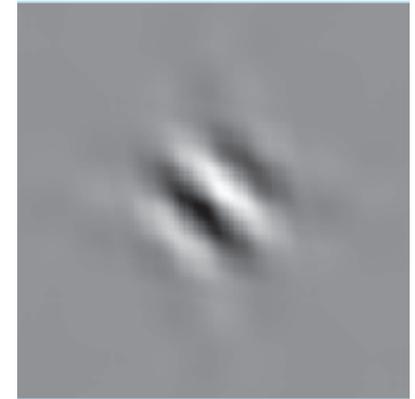
$$\psi(x, y) = \underbrace{\psi_h(x)\psi_h(y) - \psi_g(x)\psi_g(y)}_{\text{Partie réelle}} + j \underbrace{[\psi_g(x)\psi_h(y) + \psi_h(x)\psi_g(y)]}_{\text{Partie imaginaire}}$$

Partie réelle

Partie imaginaire



Enveloppe

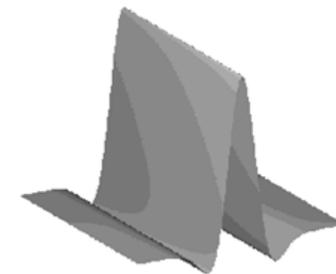


2-D : une réponse géométrique par bases fixes

Transformée géométrique à fonctions d'analyse fixes

$$r(a, b, \theta) = \int_{\mathbb{R}^2} \psi_{a,b,\theta}(\mathbf{x}) s(\mathbf{x}) d\mathbf{x}$$

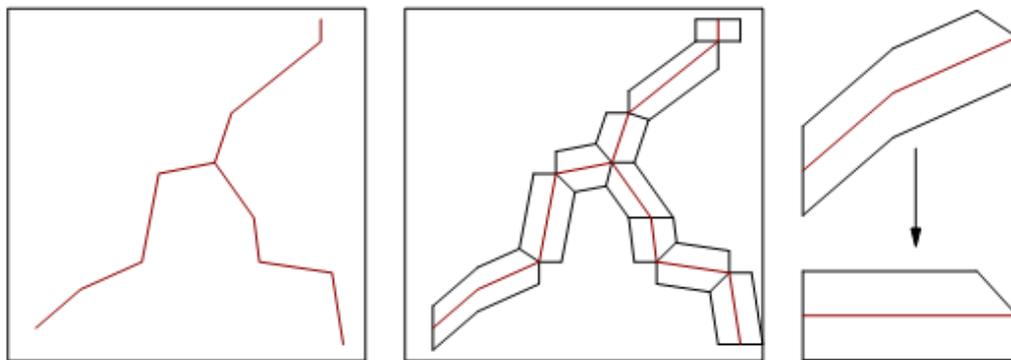
Fonction « adaptée »



➔ Ridgelet, Curvelet

2-D : une réponse géométrique par bases adaptatives

Transformées adaptatives



Bandelettes
et extensions

[Lepennec 02, Peyre05]





De nombreuses XLET
Impact dans le domaine du
tatouage / stéganalyse ?



Introduction : tatouage

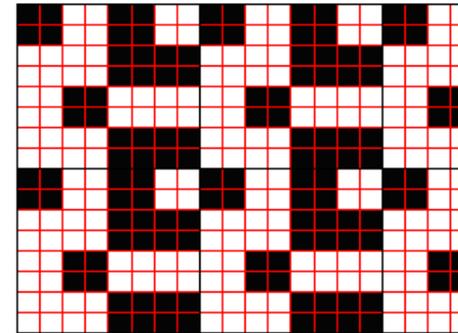
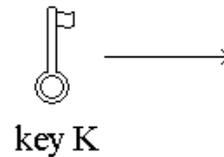
Insérer une empreinte binaire dans une image

- Applications :
Droits d'auteurs
Authentification
Indexation

...

- Attaques :

1. Bienveillantes (issues du traitement d'image : compression, filtrages, attaques géométriques ...)
2. Malveillantes



mark M

Compromis entre invisibilité, robustesse et sécurité

problème: évaluation de la dégradation d'une image

fonction de l'application



Tatouage et ondelettes

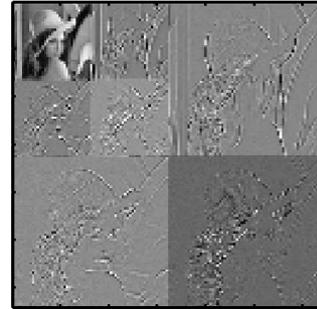
- Utiliser une stratégie spatiale : la notion de contexte locale
- Choisir une/des bandes de fréquence : la robustesse à certaines attaques, le contrôle de l'invisibilité
- Construction d'une représentation hiérarchisée
- Invariance (?)
- Offre une modélisation assez solide (Gaussienne généralisée), un lien avec le SVH



Approche naïve (Numerical Tours)

Quelques coefficients d'ondelettes

Bruit Gaussien



Insertion

$$x_i = (x_0)_i + \rho |(x_0)_i| w_i$$

Attaque

$$y = A(x)$$

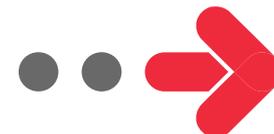
Variante

Détection

$$C(y, w) = \frac{\langle y, w \rangle}{\|y\| \|w\|}$$

Définition du seuil en fonction de la probabilité de fausse alarme

$$C(y, w) > T$$



Modification des coefficients par étalement de spectre

Sites d'insertion

$$\mathbf{y}_j = \mathbf{x}_j + \sum_{i=0}^{N_c-1} s(\mathbf{m}_j(i), \mathbf{x}_j) \mathbf{u}_i$$

Vecteur correspondant à la porteuse

Définition de la stratégie de modulation

Détection par corrélation

Sécurité, contrôle de l'invisibilité

Dimension Ondelette ?



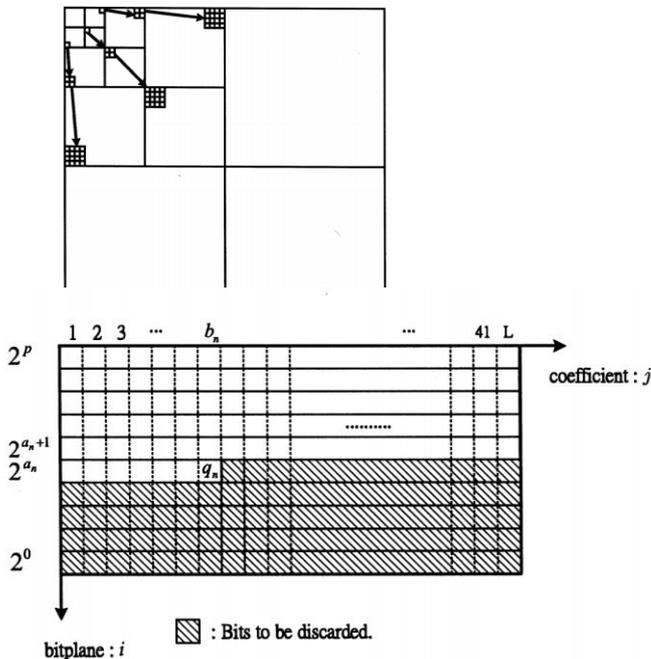
Choix des sites sur les différentes bandes avec un compromis entre robustesse et invisibilité



Modification des coefficients après groupement en arbre/en bloc

Idée : grouper les coefficients d'ondelettes et paramétrer la différence entre deux blocs/arbres

Construction JPEG2000



- Identification de deux arbres pour 1 bit message
- Quantification jusqu'à un certain niveau de plans de bits de l'un des deux arbres (selon la valeur du bit message)



Une certaine robustesse aux transformations géométriques

Dimension Ondelette ?

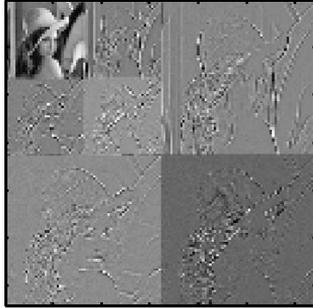
Construction hiérarchique ?

Nombre d'échelles de décomposition



QIM dans l'espace ondelette

Ensemble de coefficients d'ondelettes



$$y = Q_m(x, \Delta) = \left\lfloor \frac{x}{\Delta} \right\rfloor \Delta + (-1)^{m+1} \frac{\Delta}{4}$$

Insertion

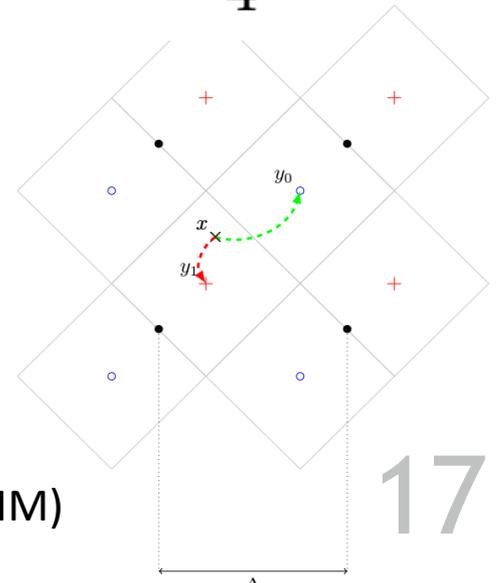
$$\Lambda_0 = \Delta \mathbb{Z}^L - \frac{\Delta}{4}$$

$$\Lambda_1 = \Delta \mathbb{Z}^L + \frac{\Delta}{4}$$

Détection

$$\hat{m} = \arg \min_{m \in \{0,1\}} \text{dist}(z, \Lambda_m),$$

$$\text{dist}(z, \Lambda) = \min_{y \in \Lambda} \|z - y\|_2$$



Adaptation de QIM

L'algorithme de Kundur → Décomposition en ondelettes

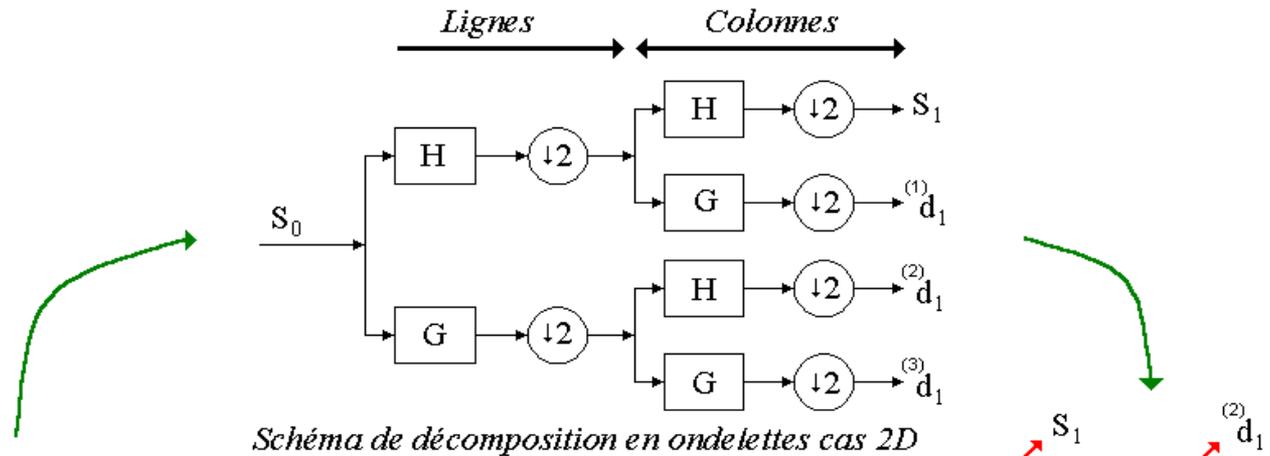


Schéma de décomposition en ondelettes cas 2D

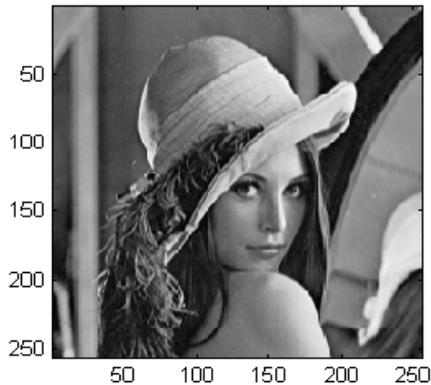
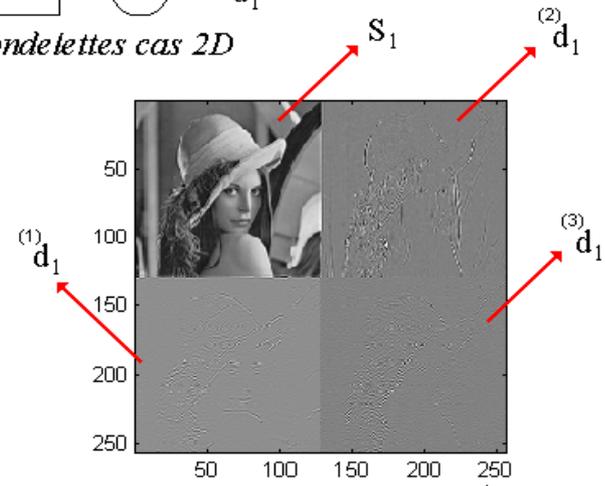
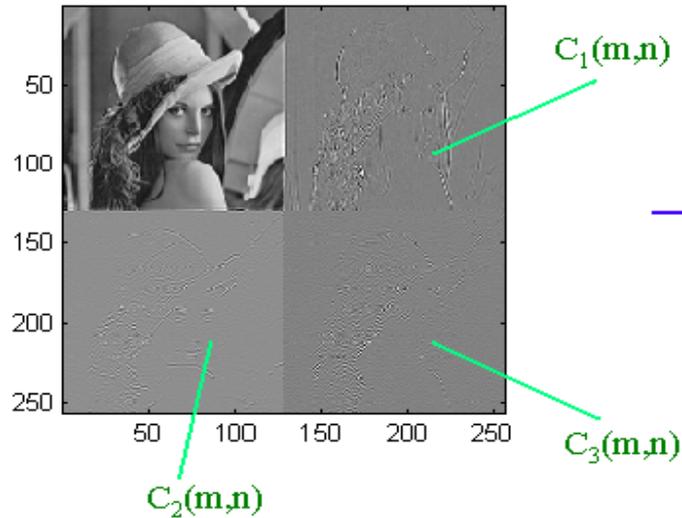


Image originale



Décomposition en ondelettes à la 1^{ère} échelle

QIM dans l'espace ondelette

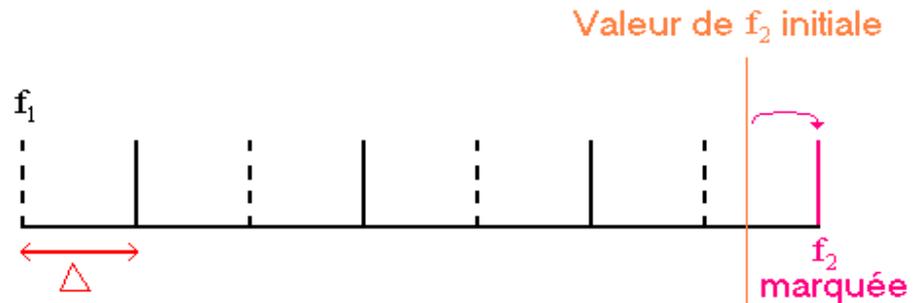


1. $\text{tri}(C_1, C_2, C_3) = (f_1, f_2, f_3)$

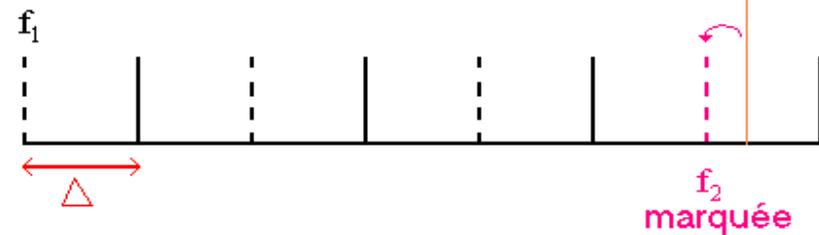
2. $\Delta = \frac{f_3 - f_1}{2Q - 1}$

Q : force du tatouage

Intégration d'un bit de valeur '1'



Intégration d'un bit de valeur '-1'



Problématique de la couleur

Image originale



Tatouage des 3 composantes R,G,B



Tatouage R



Tatouage G



Tatouage B

Invisibilité & Faible robustesse

Visibilité & Robustesse



BLEU

ROUGE

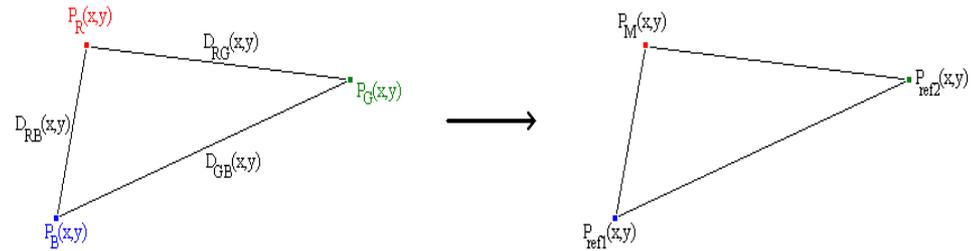
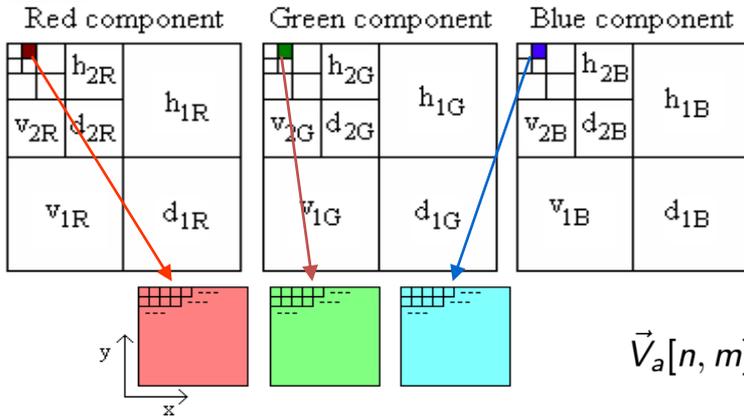
VERT



Une méthode Ad hoc

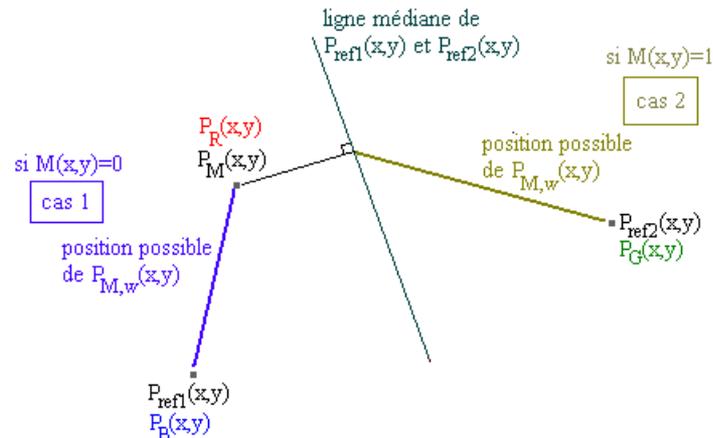
- Décomposition en ondelettes

Error correcting codes for robust color wavelet watermarking, Wadood Abdul, Philippe Carré and Philippe Gaborit



$$\vec{V}_a[n, m] = \{d_{1,L}^a[n, m], d_{2,L}^a[n, m], d_{3,L}^a[n, m]\} \quad a = \{R, G, B\}$$

- Définition d'une notion de vecteurs couleur
- Tatouage par modification vectorielle des coefficients
- Reconstruction



Invisibilité

Image originale



Image tatouée



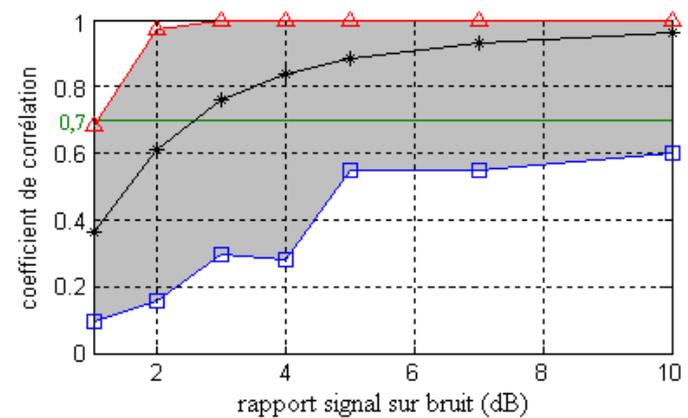
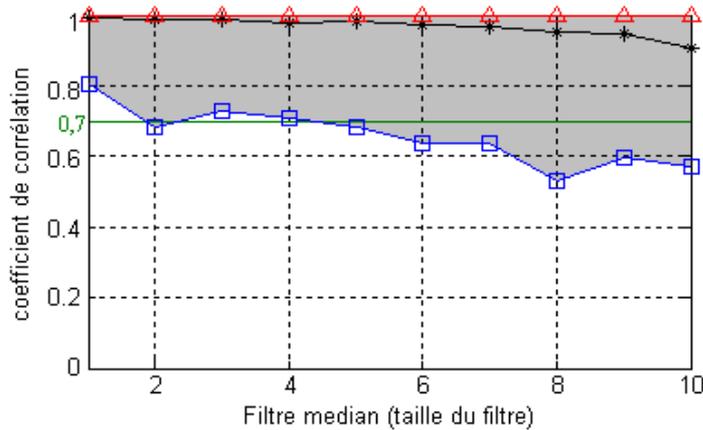
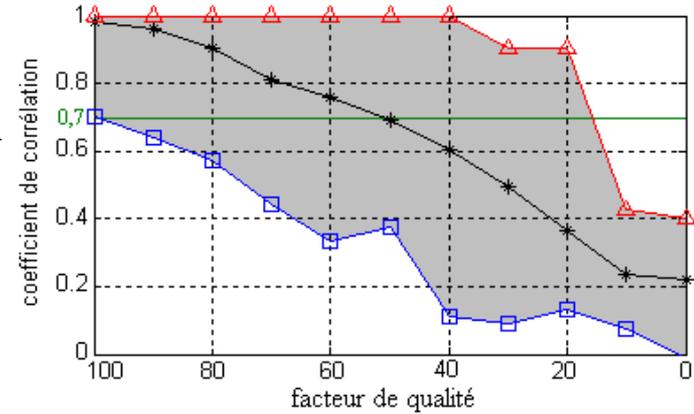
$F_R=0,6$
 $F_G=0,4$
 $F_B=0,9$



Robustesse

- △ maximum
- * moyenne
- minimum
- seuil de décision $T=0,7$

- Compression JPEG
- Ajout de bruit
- Filtrage médian



Utilisation des caractéristiques de l'espace ondelettes

1) Contrôle de la dégradation

Préciser la force de marquage maximum suivant la dégradation de l'image

Principe: deux couleurs C1 et C2 sont identiques, si dans L*a*b*:

$$\Delta E(C_1, C_2) = [(L_{C1} - L_{C2})^2 + (a_{C1} - a_{C2})^2 + (b_{C1} - b_{C2})^2]^{1/2} < 3;$$

2) Ajout d'invariance

Utilisation de la transformée en ondelettes non-décimée et code par répétition

Invisibilité optimisée

Analyse de la qualité des images tatouées

segmentation texture et
rétroaction (CIEDE2000, par
valeur moyenne)

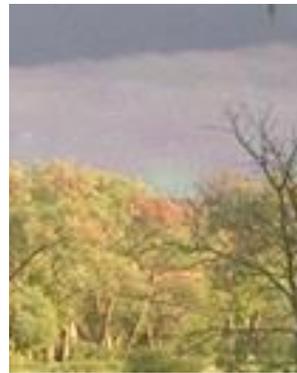


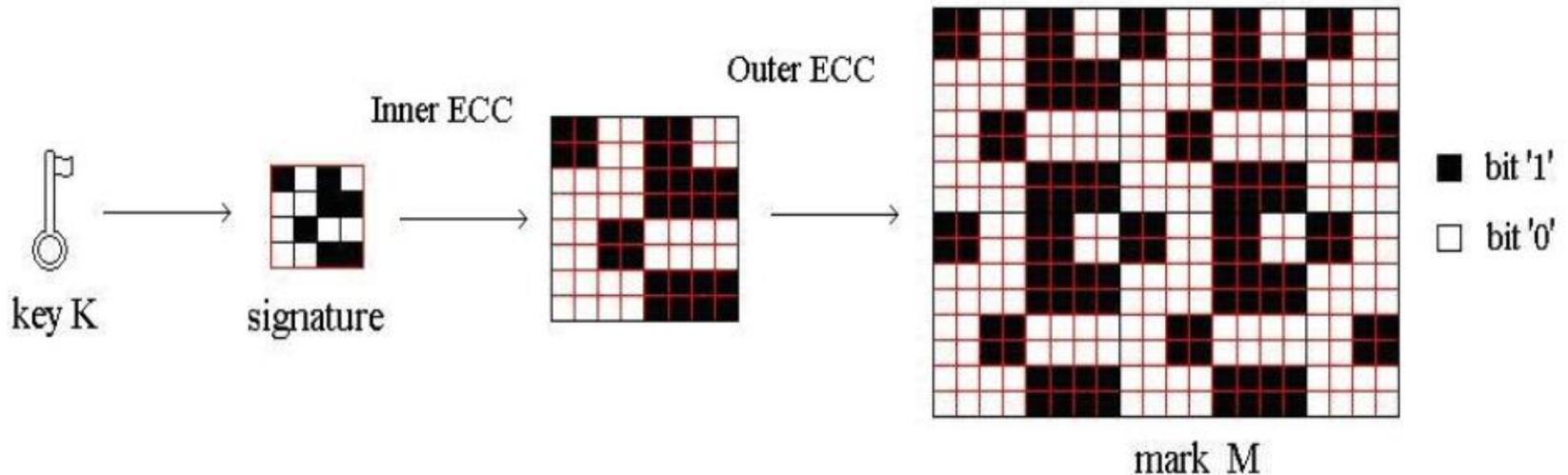
Image originale

Image tatouée
Forces de marquage
constante

Image tatouée
Segmentation et
rétroaction

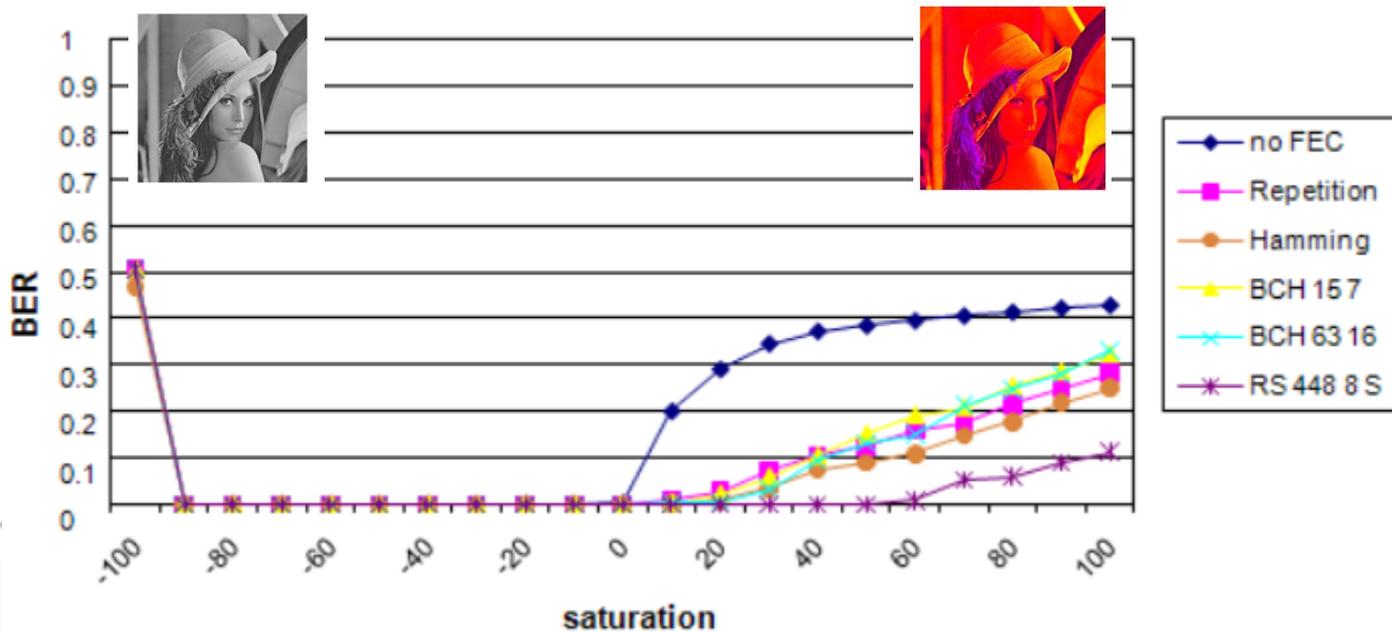
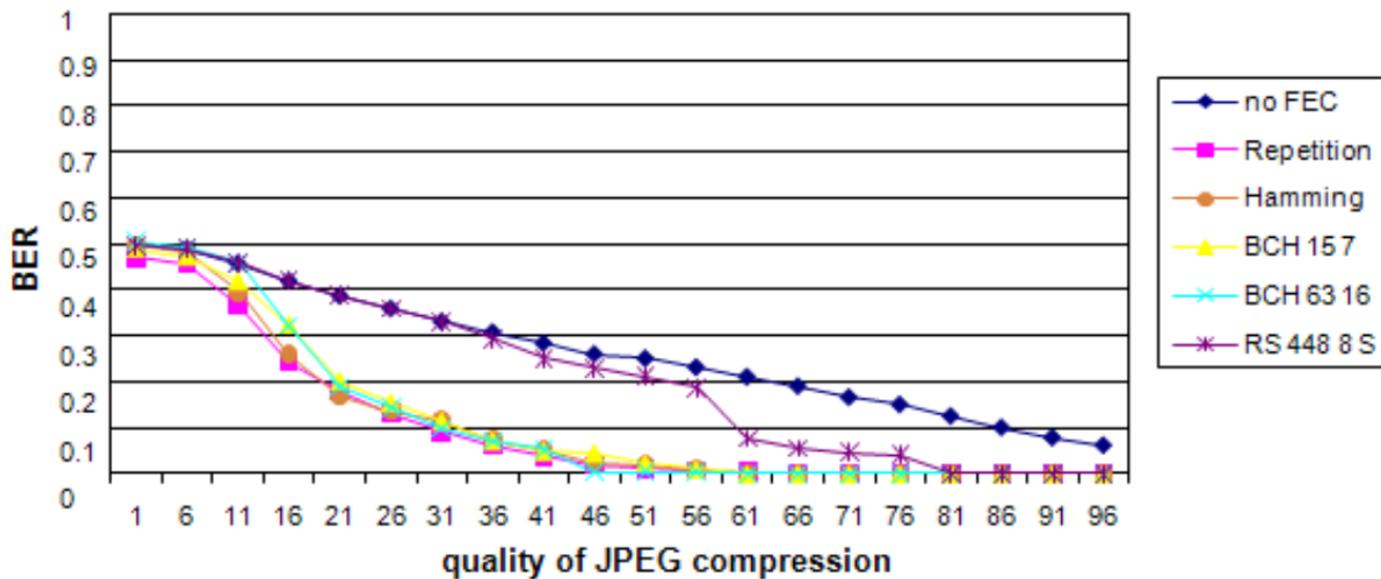


Insertion d'autres éléments : codes



- Code par répétition : erreur aléatoire
- Hamming et BCH : erreur aléatoire
- List decoding avec Reed-Solomon : erreur par plage

Attaques vs codes vs ondelettes



D'autres décompositions ?

Ondelettes non-séparables

- Tatouage par une approche par blocs
- Dynamique supérieure des coefficients aux premières échelles

A Blind Watermarking Scheme Using New Nontensor Product Wavelet Filter Banks Xinge You, Liang Du, Yiu-ming Cheung, Qihui Chen

Transformée entière

- Tatouage réversible

Reversible Image Watermarking Based on Integer-to-Integer Wavelet Transform, Jun Tian

Paramétrage des filtres par clef

PROTECTION OF WAVELET-BASED WATERMARKING SYSTEMS USING FILTER PARAMETRIZATION, Werner Dietl, Peter Meerwald, Andreas Uhl

Utilisation de la transformée CWT

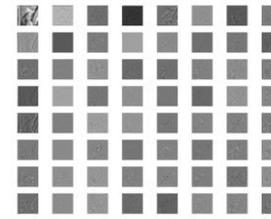
- Insertion de la « transformée » de la marque
- Utilisation d'une échelle assez grossière mais en U de YUV (Vidéo)

Imperceptible and Robust Blind Video Watermarking Using Chrominance Embedding: A Set of Approaches in the DT CWT Domain Md. Asikuzzaman, Md. Jahangir Alam, Andrew J. Lambert, Mark Richard Pickering,

D'autres décompositions ?

Ondelettes M-bandes

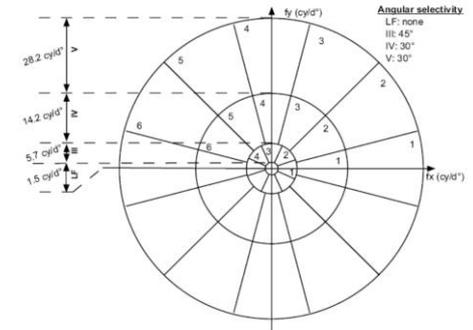
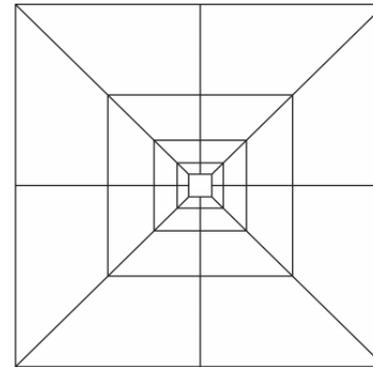
- Tatouage par une approche par blocs
- Sélection des bandes « moyennes »
- Utilisation de la décomposition EMD (comportement moyen)



Robust Image Watermarking Based on Multiband Wavelets and Empirical Mode Decomposition, Ning Bi, Qiyu Sun, Daren Huang, Zhihua Yang, and Jiwu Huang

Décomposition Contourlet

- Représentation proche du système visuel humain
- Approche (insertion) comme DWT



Intégration de masque psychovisuel, de référentiel (par ex angulaire), de choix d'échelle, de définition des blocs

Bilan pour le tatouage

- Ondelettes offrent un environnement optimisant des approches « spatiales »
- Possibilité de combiner différents éléments (code, modèle SVH, ...)
- Peu de travaux cherchant des invariances
- La reconstruction numérique peut être une attaque
- Les attaques géométriques « restent » une difficulté



Ondelette : marqueur pour la Stéga-analyse

Stéganalyse : quelques éléments

- « Targeted » Stéganalyse : Analyse d'un stégo-système particulier
- « Blind » Stéganalyse : Détection d'un message créé par un stégo-système arbitraire et inconnu

Une majorité des stéganalyseurs reposant sur de l'apprentissage machine sont considérés « Blind »

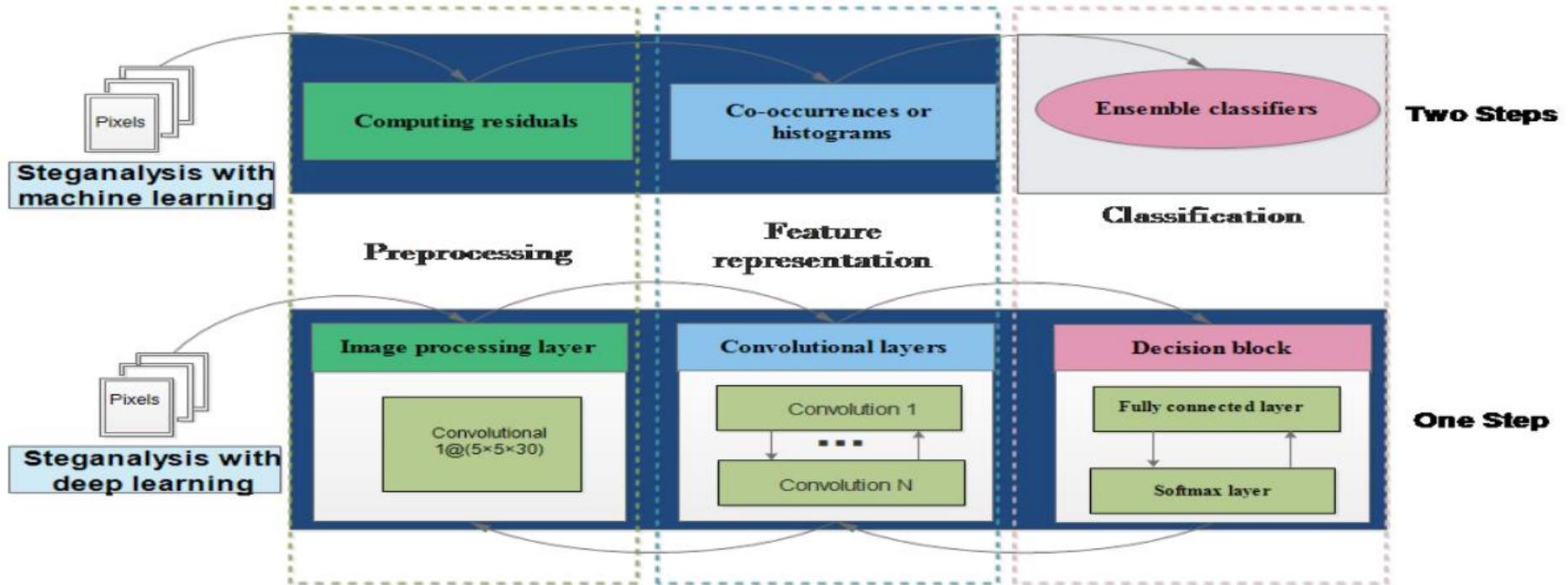


Dépendent de l'ensemble d'apprentissage : universel

Schéma classique de stéganalyse en image

1. Calcul des caractères (Rich Models)
2. Apprentissage

Stéga-analyse : Etat de l'art



M. Chaumont (Poitiers 2018)

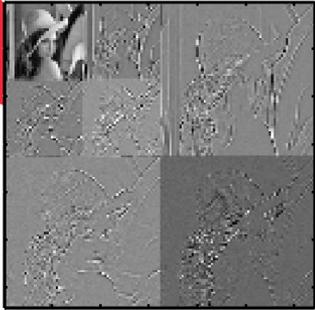
Dépendance à la base d'apprentissage, algorithme stégo

Les ondelettes en stéganalyse

- Faible présence dans la littérature
- Utilisation de la représentation « classique »
- Création de « Features » pour valider la présence d'un message

Point de départ : insuffisance de la représentation spatiale ou fréquentielle

Stratégie Classique : modélisation des coefficients



Local angular harmonic decomposition (Simoncelli) : p-dérivé orientée et notion de phase relative

Inter-échelle : log-erreur de la modélisation linéaire

$$A_2^c(x, y) = I^c(x, y) \star \left(\frac{\partial^2 G}{\partial x^2} - \frac{\partial^2 G}{\partial y^2} + 2j \frac{\partial^2 G}{\partial x \partial y} \right)$$

$$\begin{aligned} |V_i^r(x, y)| &= w_1 |V_i^r(x-1, y)| + w_2 |V_i^r(x+1, y)| \\ &+ w_3 |V_i^r(x, y-1)| + w_4 |V_i^r(x, y+1)| \\ &+ w_5 |V_i^r(x/2, y/2)| + w_6 |D_i^r(x, y)| \\ &+ w_7 |D_{i+1}^r(x/2, y/2)| + w_8 |V_i^g(x, y)| \\ &+ w_9 |V_i^b(x, y)|, \end{aligned}$$

$$\begin{aligned} s_{p,q}^{c_1, c_2}(x, y) &= \sqrt{|A_q^{c_1}(x, y)| |A_p^{c_2}(x, y)|} \\ &\times \exp(j\phi_{p,q}^{c_1, c_2}(x, y)), \end{aligned}$$

$$\vec{p} = \log(\vec{v}) - \log(|Q\vec{w}|)$$

Coefficients de la DWT

Modélisation : moment d'ordre n (mean, variance, skewness, kurtosis ...)

$$m_n = EX^n = \int_{-\infty}^{\infty} p(x) x^n dx.$$

Steganalysis Using Higher-Order Image Statistics
Siwei Lyu, and Hany Farid

Classifieur (SVM)

Stratégie Classique : modélisation des coefficients

Estimation du signal Stégo

Blind Statistical Steganalysis of Additive Steganography Using Wavelet Higher Order Statistics Taras Holotyak, Jessica Fridrich, Sviatoslav Voloshynovskiy

- « Débruitage » dans le domaine des ondelettes pour estimer le signal Cover
- Modélisation du signal Stégo par le calcul de la distribution des coefficients de la première échelle du signal de différence (utilisation des moments)
- Classification par « Fisher Linear Discrimination »

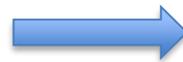
Modélisation « directe » des coefficients de la décomposition en ondelettes

Effective Steganalysis Based on Statistical Moments of Wavelet Characteristic Function Yun Q. Shi et al.

Moment de la fonction caractéristique de la première échelle

Bilan Ondelettes

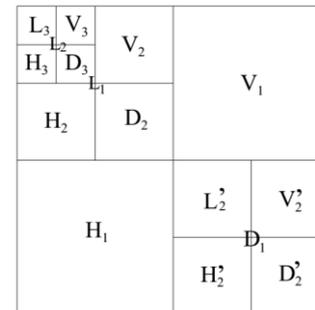
- Modélisation des coefficients du bruit (Signal Stego)
- Modélisation de l'erreur de prédiction
- Moments empirique de la PDF (estimés à partir des coefficients) ou de la fonction caractéristique (estimée à partir de l'histogramme)



Classifieur

Bilan

Optimized Feature Extraction for Learning-Based Image Steganalysis Ying Wang, Pierre Moulin



Utilisation du banc de filtres, pas de l'outil d'analyse

Wavestega : un projet Multidisciplinaire

- Signal, Représentations de type ondelettes,
- Stéga-analyse,
- Analyse Multifractale,
- Tests statistique,
- Développement numérique.

Projet Imagin



P. Carré F. Enikeeva



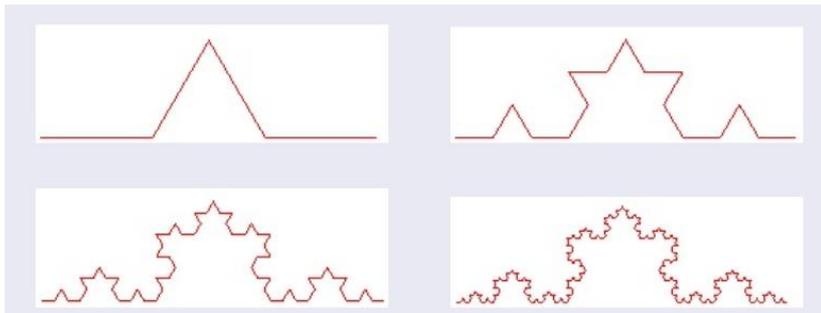
M. Clausel L. Navarro



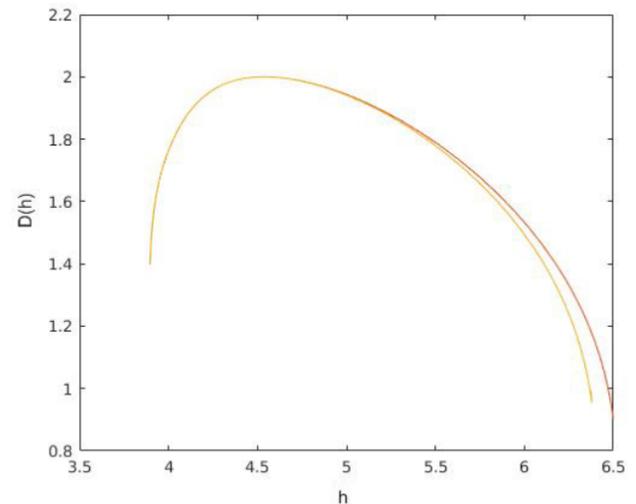
proposition : création d'une signature et d'un test associé

L'insertion de ce message va venir perturber la structure de l'image

Mesure : spectre multifractal



Sans apprentissage



Mesure : régularité locale

- Exposant de Hölder de la fonction X en t_0 (cas où la fonction est nulle part différentiable)

$$h_X(t_0) = \sup \{ \alpha \geq 0 : |X(t) - X(t_0)| \leq C|t - t_0|^\alpha \}.$$

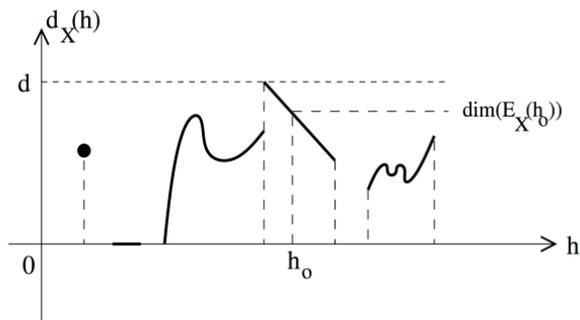
$h_X(t_0)$ des informations significatives sur le comportement local de X autour de t_0 : plus il est faible plus le "comportement" de la fonction X est irrégulier.

Le spectre multifractale

- Utilisation des ensembles iso-Hölder $E_X(h) = \{t, h_X(t) = h\}$ afin de décrire la diversité des comportements locaux de X .
- Spectre multifractal, associe à $h \in (0, 1)$, la “taille” des ensembles de points (au sens de la dimension de Hausdorff) d'exposant de Hölder h

$$D_X : h \mapsto \dim_H(E_X(h))$$

Le spectre multifractal : la distribution géométrique des singularités de X .



Estimation numérique du spectre des singularités

Comportement en loi de puissance en fonction de l'échelle analysée $S_X(a, q) := \frac{1}{n_a} \sum_k |T_X(a, k)|^q = C_q * (a)^{\zeta(q)}$ avec $T_X(a, k)$ coefficients multirésolution, Fonction d'échelle

le spectre multifractal vérifie la relation (spectre de Legendre)

$$D_X(h) = \inf_{q \in \mathbb{R}} \{qh - \zeta(q) + d\}$$

Hölder d=1,2

avec $\zeta(q)$ une mesure globale appelée fonction d'échelle.
Spectre multifractal calculé à partir de la fonction d'échelle





Ondelette et Multifractale

Les ondelettes : un outil pour l'étude multifractale.

- Etude locale : régularité locale (Hölder) connectée aux coefficients d'ondelettes,
- Régularité liée à la structure inter-échelle du signal (caractérisée par la décroissance des coefficients d'ondelettes à travers les échelles).

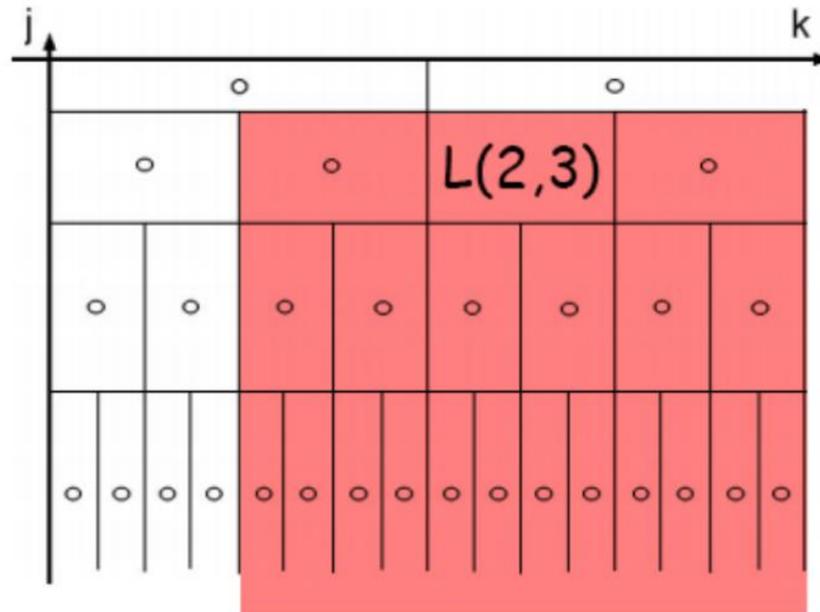
⇒ Une généralisation de l'outil : les Wavelet Leaders

- Lien connu entre **Wavelet Leaders** et exposant de Hölder.
- Quelques références :
 - S. Jaffard, B. Lashermes, and P. Abry. Wavelet leaders in multifractal analysis. In *Wavelet Analysis and Applications*, Birkhkauser Verlag, 2006.
 - H. Wendt, S. G. Roux, S. Jaffard, and P. Abry. Wavelet leaders and bootstrap for multifractal Analysis of images. *Signal Process.*, June 2009

Wavelet Leader

On note $L_X(j; k)$ les wavelet leaders associés au signal X à l'échelle d'observation j et à la position k :

$$L_X(j; k) = \max_{\lambda' \subset 3\lambda_{j,k}} |c_X(\lambda')|$$



Estimation du spectre

En pratique : calcul de la courbe paramétrée $(h(q), D_X(q))$.

Point de départ

- Fonction d'échelle et Wavelet Leader

Fonction d'échelle

$$S_X(j, q) := \frac{1}{n_j} \sum_{k=1}^{n_j} L_X^q(j, k) \approx C_q * (2^j)^{\zeta^L(q)}$$

- Rappel : spectre multifractal calculé à partir de la fonction d'échelle

$$D_X(h) = \min_{q \neq 0} \{qh - \zeta^L(q) + d\}$$

Estimation du spectre

Régression linéaire

- On estime

$$\tilde{h}(q) = \sum_{j=j_1}^{j_2} w_j \frac{\sum_{k=1}^{n_j} L^q(j, k) \log_2 L(j, k)}{\sum_{k=1}^{n_j} L^q(j, k)}$$

avec w_j coefficients de régression linéaire (de la dérivé de la fonction d'échelle vs l'échelle).

- on estime

$$D_X(q) = d + \sum_{j=j_1}^{j_2} w_j \frac{\sum_{k=1}^{n_j} L^q(j, k) \log_2 L^q(j, k) - L^q(j, k) \log_2 \sum_{k=1}^{n_j} L^q(j, k)}{\sum_{k=1}^{n_j} L^q(j, k)} + \log_2(n_j)$$

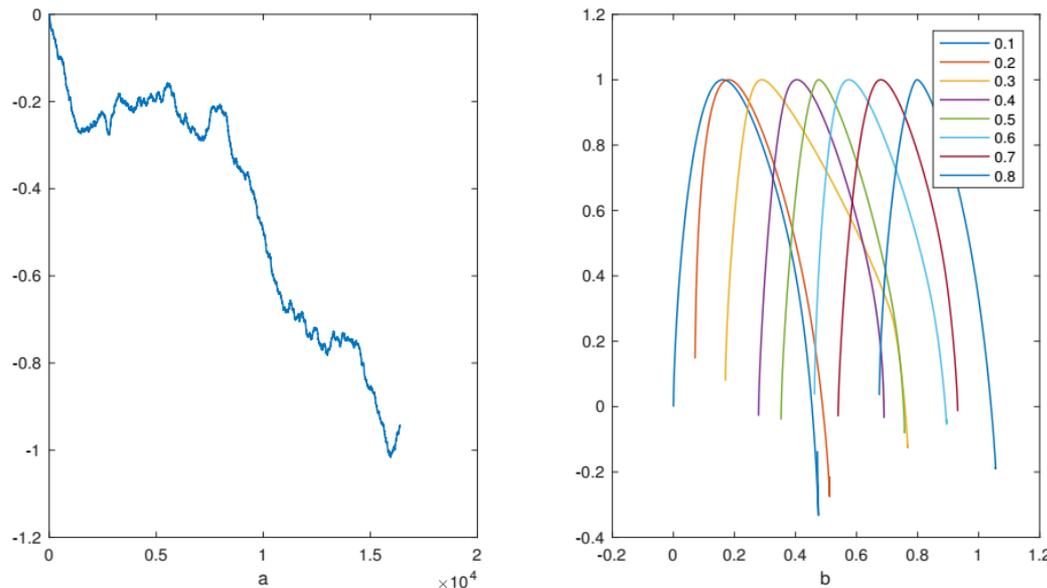
avec w_j coefficients de régression linéaire (de l'expression de $D(q)$ vs l'échelle).

⇒ On obtient la courbe paramétrique $(h(q), D_X(q))$

Exemple de spectre

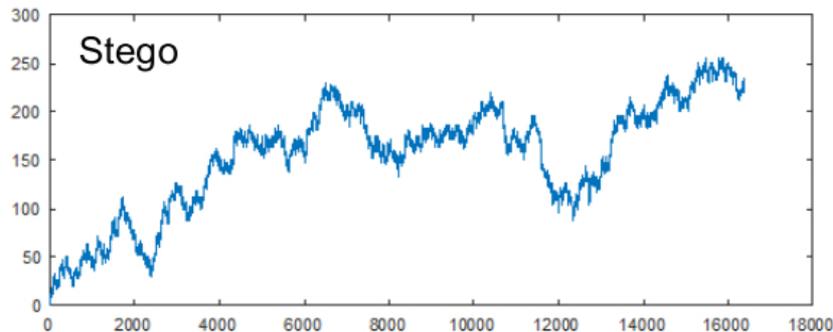
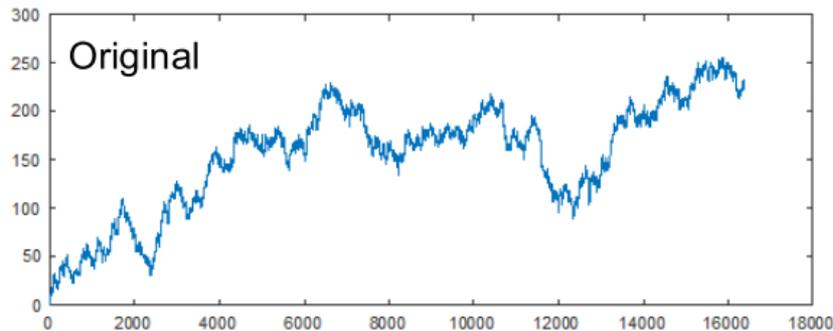
Le fBm (Fractional Brownian motion) $X(t) := W_H(t)$, avec H le paramètre d'autosimilarité, $t \in \mathbb{R}^+$ est un processus monofractal : $\forall t \in \mathbb{R}^+$ (Un spectre qui doit se réduire à 1 point).

$$\alpha(t) = H \implies d_X(h) = \delta_H(h).$$



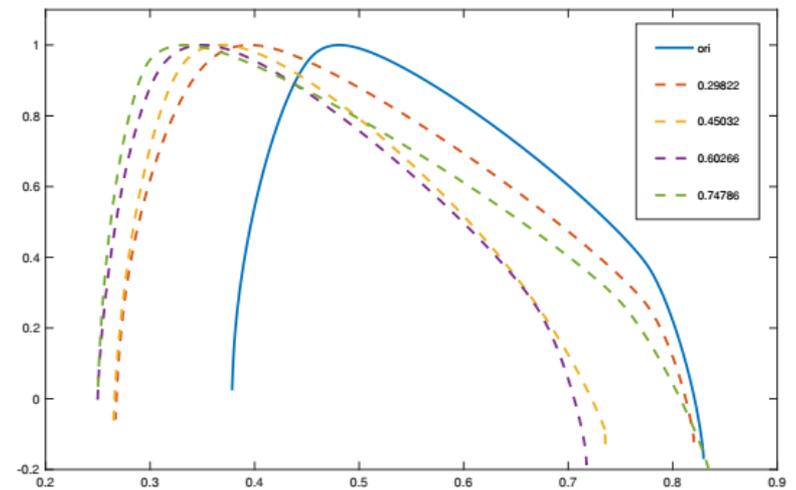
Exemple de perturbation

- Donnée : le fBm $W_H(t)$, $H = 1/2$
- Modification par un algorithme de stéganographie LSB
- Contexte : 2 LSB sur 8 bits avec un taux d'insertion $p \in (0, 1)$



Irrégulier

Régulier

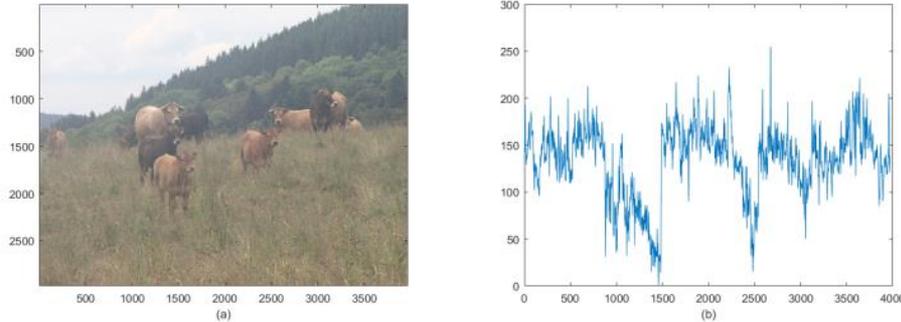


Le spectre du fBm $W_{0.5}(t)$ stéganographié pour différentes valeurs de p , la courbe bleue est le spectre original ($p = 0$)

Insertion invisible d'un message $p = 0.4$

Donnée réelle

- Donnée : une ligne extraite d'une image :



- Contexte : 2 LSB sur 8 bits avec un taux d'insertion $p \in (0, 1)$

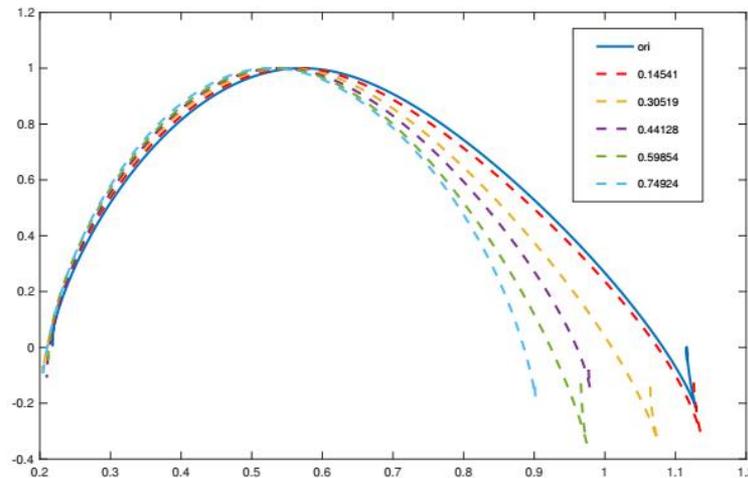
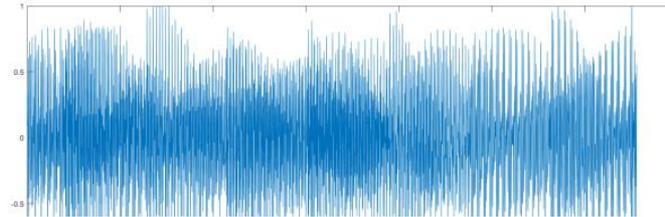
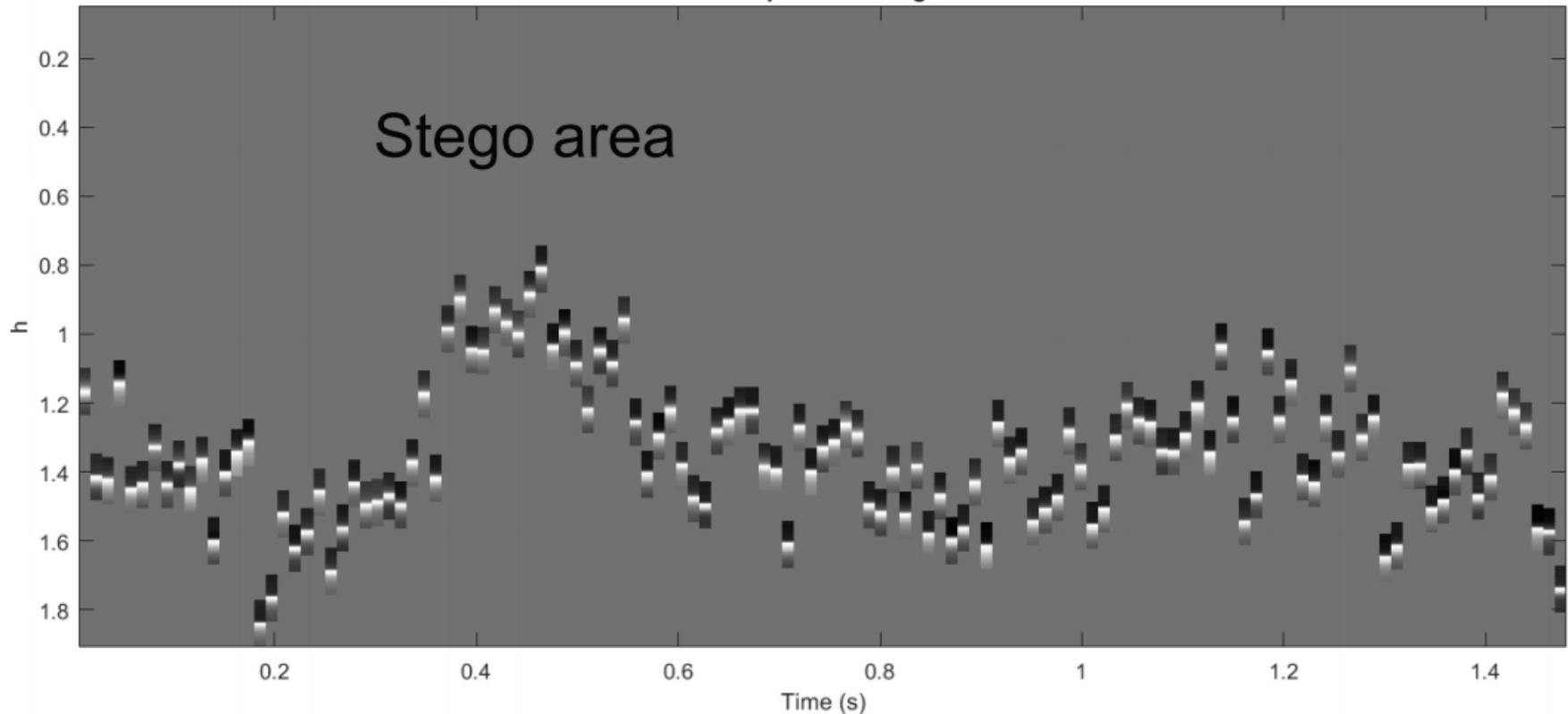


Figure – Le spectre de la ligne d'image pour différentes valeurs de p . La courbe bleue est le spectre original

- Enregistrement d'une note sur le dynamique de 8 bits.
- Contexte : 2 LSB sur 8 bits avec un taux d'insertion $p \in (0, 1)$



Multifractal spectrum along the time



Prise de décision

- Conjecture : insertion d'un message change la régularité d'une image dans les espaces L^q et donc sa fonction d'échelle/spectre multifractal
- Test d'hypothèses sur la fonction d'échelle :

$$H_0 : \forall q, \zeta(q) \geq \zeta_0(q) \quad \text{contre} \quad H_1 : \forall q, \zeta(q) < \zeta_0(q)$$

où $\zeta_0(q)$ est la fonction d'échelle de l'image vierge.

- Étape préliminaire : être capable à q fixé de faire un test de régularité de l'image dans L^q