Estimateurs statistiques et décompositions directionnelles

Laurent Duval, IFP Energies nouvelles

avec A. Benazza-Benyahia, C. Chaux, J. Gauthier, J.-C. Pesquet

2013

Introduction : plan

- Estimateurs statistiques standard
- Décompositions directionnelles multiéchelle
- Applications au débruitage
- Conclusions, pointeurs

Applications : motivations initiales

- Traitement de données volumineuses
  - variété des types (image, sismique), échantillonnage hétérogènes, anisotropie
- Besoins
  - dév. de classes flexibles de transformations (1,2,3-D)
  - directionnalité, faible redondance, support spatial localisé
  - (idée de représentations complexes)
    1. bancs de filtres M-bandes complexes suréch. (OSFB)
    2. ondelettes M-bandes en arbre dual (DTT)
- Débruitage statistique
  - minimum d'hypothèses a priori (sur le bruit)
  - lemme de Stein et SURE (Stein Unbiased Risk Estimation)
    1. combinaison linéaire de seuillages (SURE-LET)
    2. estimateur paramétrique "générique"

Applications : méthodologie

Deux approches :

- une low-cost, avec comparaison analyse/synthèse
  - trame non-ajustée (effort sur les filtres)
  - redondance très réduite $(< 2)^d$
  - seuillage scalaire par groupes homologues
- une plus évoluée
  - trame potentiellement ajustée
  - redondance faible (2)
  - multicomposante (couleur), persistance multi-échelle, étalement du bruit

Applications : approche 1

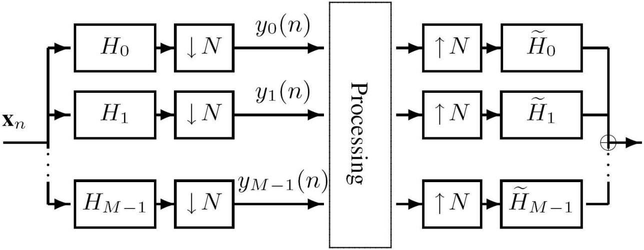
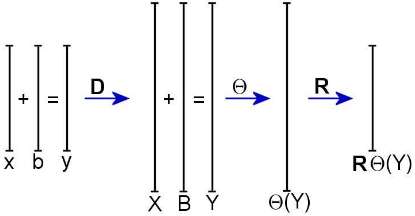


Figure : Banc de filtres redondant complexe (cf. STFT)

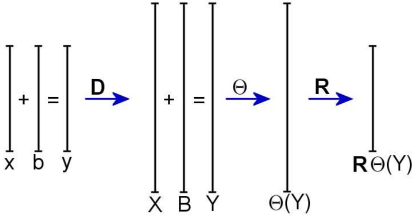
Applications : problème de débruitage

Notation :



Applications : problème de débruitage

Notation :



Objectifs :

- Trouver un estimateur $\mathbf{G}(\mathbf{y})$ de $\mathbf{x}$ minimisant l'erreur quadratique moyenne : $E\left(\|\mathbf{G}(\mathbf{y})-\mathbf{x}\|^2\right)$.
- Trouver un estimateur $\mathbf{F}(\mathbf{Y})$ of $\mathbf{X}$ minimisant l'erreur quadratique moyenne : $E\left(\|\mathbf{F}(\mathbf{Y})-\mathbf{X}\|^2\right)$.

Applications : principe de Stein

Lemme (*[Stein 1981]*)

Soit $\boldsymbol{G} : \mathbb{R}^L \to \mathbb{R}^L$ *une fonction continue, différentiable presque partout, telle que :*

1. $\mathsf{E}\left(\|\boldsymbol{G}(\boldsymbol{y})\|^2\right) < \infty$,
2. $\mathsf{E}\left(\left\|\frac{\partial \boldsymbol{G}(\boldsymbol{y})}{\partial \boldsymbol{y}^\top}\right\|_F\right) < \infty$,
3. $\forall \boldsymbol{z} \in \mathbb{R}^L, \quad \lim_{\|\boldsymbol{t}\| \to +\infty} \boldsymbol{G}(\boldsymbol{t}) \exp\left(-\frac{(\boldsymbol{t}-\boldsymbol{z})^\top (\boldsymbol{\Gamma}^{(\boldsymbol{b})})^{-1}(\boldsymbol{t}-\boldsymbol{z})}{2}\right) = \boldsymbol{0}$,

alors :

$$\mathsf{E}(\boldsymbol{G}(\boldsymbol{y})\boldsymbol{x}^\top) = \mathsf{E}(\boldsymbol{G}(\boldsymbol{y})\boldsymbol{y}^\top) - \mathsf{E}\left(\frac{\partial \boldsymbol{G}(\boldsymbol{y})}{\partial \boldsymbol{y}^\top}\right)\boldsymbol{\Gamma}^{(\boldsymbol{b})}.$$

Applications : SURE-LET

Principe :

- Considérer plusieurs fonctions (de seuillage) : $\mathbf{F}_k$ or $\mathbf{G}_k$
- Combinaison linéaire de différentes fonctions : $\mathbf{F} = \sum_k a_k \mathbf{F}_k$ or $\mathbf{G} = \sum_k a_k \mathbf{G}_k$
- Pour chaque estimateur de Stein de risque
  $\rightarrow$ calculer les paramètres a_k (par minimisation)
- *[Pesquet, Leporini 1997], [Raphan, Simoncelli 2006], [Luisier, Blu 2007]*

Applications : SURE-LET

Principe :

- Considérer plusieurs fonctions (de seuillage) : $\mathbf{F}_k$ or $\mathbf{G}_k$
- Combinaison linéaire de différentes fonctions : $\mathbf{F} = \sum_k a_k \mathbf{F}_k$ or $\mathbf{G} = \sum_k a_k \mathbf{G}_k$
- Pour chaque estimateur de Stein de risque
  $\rightarrow$ calculer les paramètres a_k (par minimisation)
- *[Pesquet, Leporini 1997], [Raphan, Simoncelli 2006], [Luisier, Blu 2007]*

Deux applications possibles du principe de Stein :

- Après reconstruction (sur les échantillons) $\rightarrow$ fb-sure-let-s.
- Après décomposition (sur les coefficients) $\rightarrow$ fb-sure-let-c.

Applications : SURE-LET

Principe :

- Considérer plusieurs fonctions (de seuillage) : $\mathbf{F}_k$ or $\mathbf{G}_k$
- Combinaison linéaire de différentes fonctions : $\mathbf{F} = \sum_k a_k \mathbf{F}_k$ or $\mathbf{G} = \sum_k a_k \mathbf{G}_k$
- Pour chaque estimateur de Stein de risque
  $\rightarrow$ calculer les paramètres a_k (par minimisation)
- *[Pesquet, Leporini 1997]*, *[Raphan, Simoncelli 2006]*, *[Luisier, Blu 2007]*

Deux applications possibles du principe de Stein :

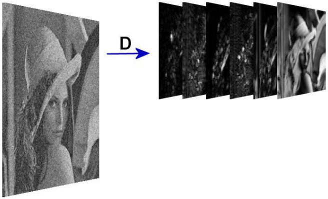
- Après reconstruction (sur les échantillons) $\rightarrow$ fb-sure-let-s.
- Après décomposition (sur les coefficients) $\rightarrow$ fb-sure-let-c.

Dans le cas orthogonal $\rightarrow$ les 2 approches sont identiques.

Applications : SURE et échantillons

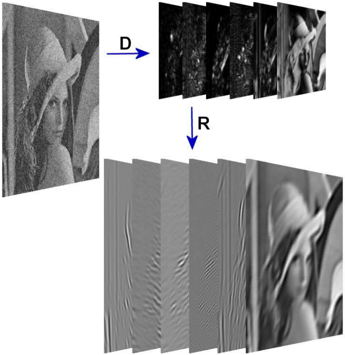
Image bruitée.

Applications : SURE et échantillons



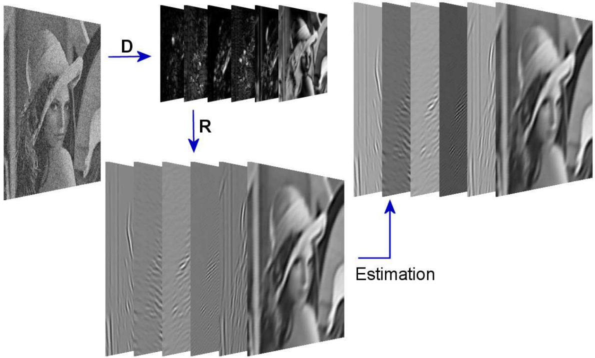
Décomposition.

Applications : SURE et échantillons



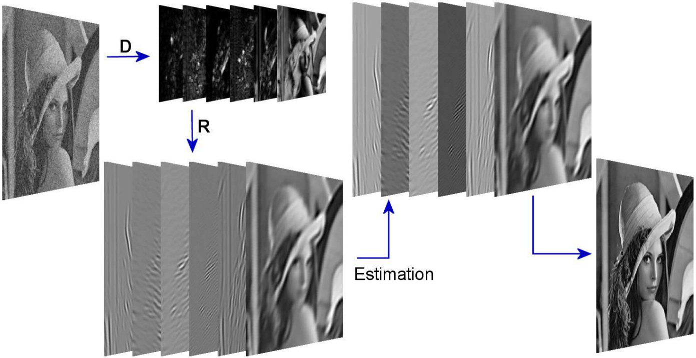
Reconstruction avec les fonctions de seuillage (fonctions $\mathbf{G}_k$).

Applications : SURE et échantillons



Minimisation de l'estimateur de risque de Stein calculé sur les images reconstruites $\rightarrow$ estimation des paramètres a_k.

Applications : SURE et échantillons



Combination linéaire → image débruitée.

Applications : SURE et échantillons

Estimateur :

$$\varepsilon_{\mathrm{S}} = \|\mathbf{G}(\mathbf{y}) - \mathbf{y}\|^2 + 2\sigma^2 \mathrm{div}(\mathbf{G}(\mathbf{y})) - L\sigma^2$$

estimateur non-biaisé de $\mathsf{E}\left(\|\mathbf{G}(\mathbf{y}) - \mathbf{x}\|^2\right)$.

Divergence (cas complexe) :

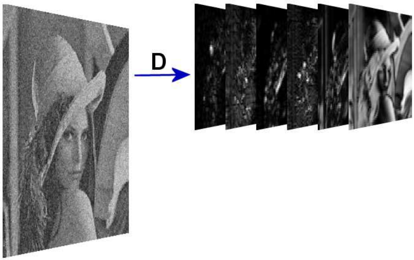
$$\mathrm{div}(\mathbf{G}(\mathbf{y})) = \mathrm{diag}(\boldsymbol{D}^R\boldsymbol{R})^\top \underline{\Theta}^{(1)}(\mathbf{Y}^R, \mathbf{Y}^I) + \mathrm{diag}(\boldsymbol{D}^I\boldsymbol{R})^\top \underline{\Theta}^{(2)}(\mathbf{Y}^R, \mathbf{Y}^I),$$

avec $.^R$: partie réelle et $.^I$: partie imaginaire.

Applications : SURE et coefficients

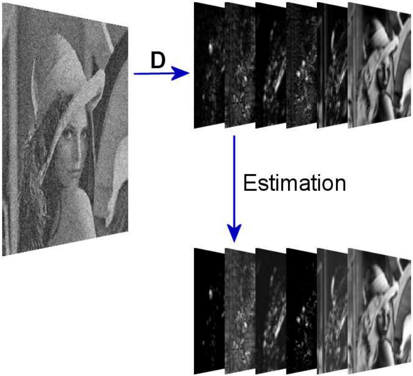
Image bruitée.

Applications : SURE et coefficients



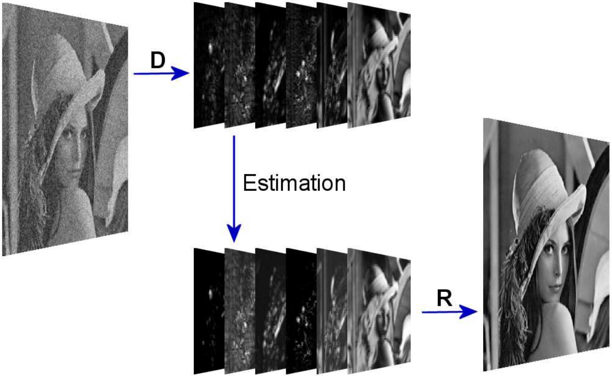
Décomposition et seuillage (fonctions $\mathbf{F}_k$).

Applications : SURE et coefficients



Minimisation de l'estimateur de risque de Stein $\rightarrow$ calcul des paramètres a_k.

Applications : SURE et coefficients



Combination linéaire et reconstruction → image débruitée.

Applications : SURE et coefficients

Estimateur :

$$\varepsilon_{\mathrm{C}} = \sum_{j=1}^{L'} \left(\left\| \vec{\Theta}_j(\mathbf{Y}_j) - \mathbf{Y}_j \right\|^2 + 2\mathrm{Tr}\left(\frac{\partial \vec{\Theta}_j(\mathbf{Y}_j)}{\partial \mathbf{Y}_j^\top} \mathbf{\Gamma}_j \right) - \mathrm{Tr}(\mathbf{\Gamma}_j) \right)$$

estimateur non-biaisé $\mathsf{E}\left(\|\mathbf{F}(\mathbf{Y}) - \mathbf{X}\|^2 \right)$.

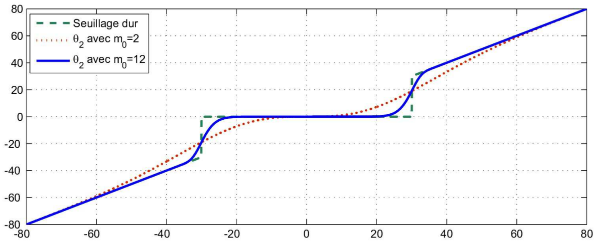
Applications : fonctions "de seuillage"

En chaque coefficient :

$$\theta_1(x,y) = x + \imath y, \quad \forall (x,y) \in \mathbb{R}^2$$

$$\theta_2(x,y) = (x + \imath y)\left(1 - e^{-\frac{\left(x^2+y^2\right)^{\frac{m_0}{2}}}{(\beta\sigma)^{m_0}}}\right), \quad \forall (x,y) \in \mathbb{R}^2,$$

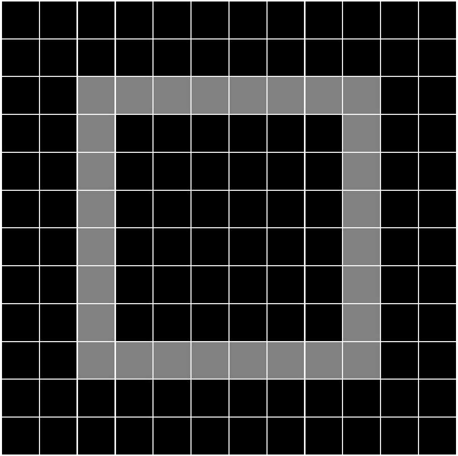
avec $m_0 \in \mathbb{R}_+^*$ et $\beta \in \mathbb{R}_+^*$.



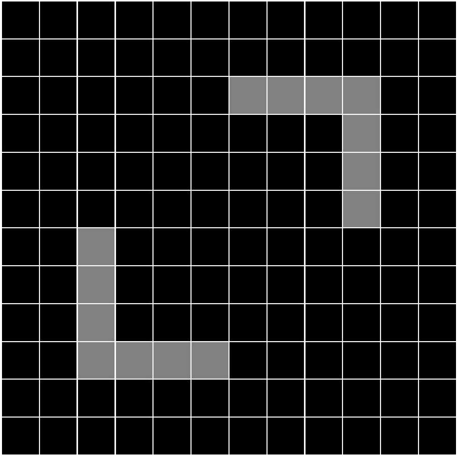
Applications : fonctions "de seuillage"

Applications : fonctions "de seuillage"

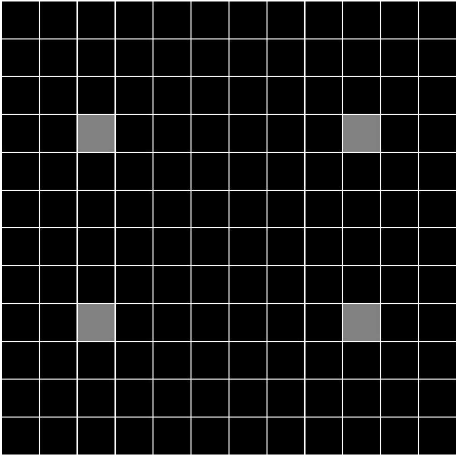
Applications : fonctions “de seuillage”



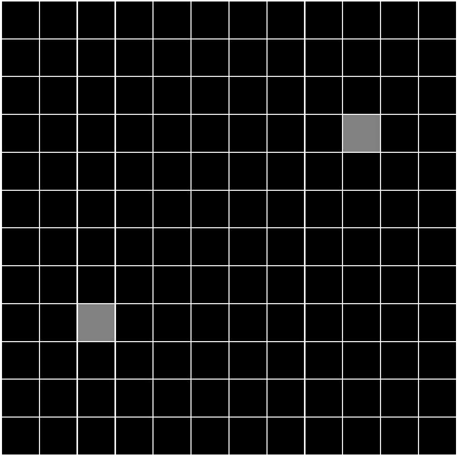
Applications : fonctions “de seuillage”



Applications : fonctions “de seuillage”



Applications : fonctions “de seuillage”



Applications : FB-SURE vs oracle

	Lena 256×256	*Lena* 512×512
Image bruitée	22.1	22.1
fb-sure-let-s	30.0	31.3
Oracle-S	30.2	31.3
fb-sure-let-c	29.2	30.5
Oracle-C	29.6	30.7

Table : Débruitage (PSNR) avec fb-sure-let-s et fb-sure-let-c ainsi que leurs oracles respectifs pour Lena avec un bruit de $\sigma = 20$.

Applications : FB-SURE vs redondance

Redondance k'	$5/4$	$3/2$	$7/4$	2
fb-sure-let-c	16.7	24.3	25.5	25.9
fb-sure-let-s	26.0	26.7	26.9	27.0
Différence	9.3	2.4	1.4	1.1

Redondance k'	$9/4$	$5/2$	$11/4$	3
fb-sure-let-c	26.0	25.7	25.8	27.1
fb-sure-let-s	27.2	26.8	26.9	27.4
Différence	1.2	1.1	1.1	0.3

Table : PSNR (en dB) après reconstruction pour les deux méthodes, pour Lena ($\sigma = 30$), avec $N = 8$, $k = 3$ et différents k'.

Applications : FB-SURE vs redondance

Redondance k'	5/4	3/2	7/4	2
fb-sure-let-c	16.7	24.3	25.5	25.9
fb-sure-let-s	26.0	26.7	26.9	27.0
Différence	9.3	2.4	1.4	1.1

Redondance k'	9/4	5/2	11/4	3
fb-sure-let-c	26.0	25.7	25.8	27.1
fb-sure-let-s	27.2	26.8	26.9	27.4
Différence	1.2	1.1	1.1	0.3

Table : PSNR (en dB) après reconstruction pour les deux méthodes, pour Lena ($\sigma = 30$), avec $N = 8$, $k = 3$ et différents k'.

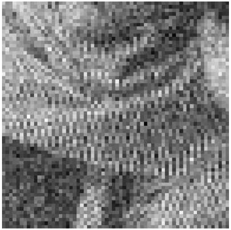
Ratio temps de calcul fb-sure-let-cvs fb-sure-let-s: 1/12

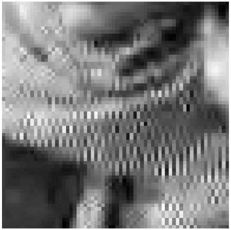
Applications : FB-SURE vs. autres méthodes

σ	10	20	30	40	50
Image bruitée	28.1	22.1	18.6	16.1	14.2
Lena					
Curvelets	33.0	29.3	27.3	25.9	24.8
SureShrink CS	32.6	28.9	26.4	25.5	24.0
BiShrink	32.8	29.0	26.9	25.5	24.6
UWT sure-let	33.0	29.4	27.4	26.1	25.2
fb-sure-let-c	33.0	29.3	27.1	25.5	24.3
fb-sure-let-s	**33.8**	**30.0**	**28.0**	**26.4**	**25.3**
Barbara					
Curvelets	31.2	27.2	25.1	23.6	22.6
SureShrink CS	30.7	26.5	24.1	21.9	21.4
BiShrink	30.1	26.6	24.4	23.1	22.2
UWT sure-let	30.8	26.4	24.3	23.2	22.4
fb-sure-let-c	31.6	27.6	25.4	23.7	22.3
fb-sure-let-s	**32.2**	**28.3**	**26.2**	**24.7**	**23.6**

Applications : résultats visuels


Original


Bruitée


BiShrink

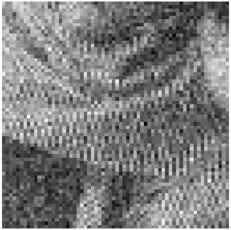

Curvelets

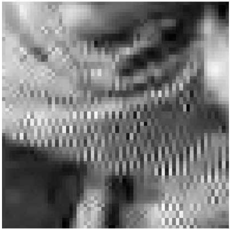

UWT sure-let


fb-sure-let-c

Applications : résultats visuels


Original


Bruitée

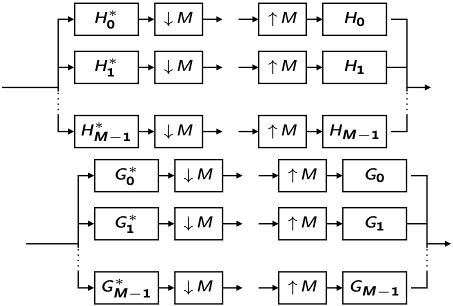

BiShrink


Curvelets


UWT sure-let


fb-sure-let-s

Applications : approche 2



Objectif

$\forall M \geq 2$ et $M \in \mathbb{N}$, construire une décomposition en ondelettes M-bandes orthogonale "duale" associée à la fonction d'échelle ψ_0^{H} et aux ondelettes mères ψ_m^{H}, où

$$\forall m \in \{1, \ldots, M-1\}, \quad \widehat{\psi}_m^{\mathrm{H}}(\omega) = -\imath \operatorname{sign}(\omega) \widehat{\psi}_m(\omega)$$

G_m, $m \in \{0, \ldots, M-1\}$: filtres associés à la décomposition duale.

Applications : M-bandes en arbre dual

- Décomposition en ondelettes M-bandes orthogonales of $\mathrm{L}^2(\mathbb{R})$
  - deux ensembles de filtres : $(H_m)_{0\leq m<M}$ and $(G_m)_{0\leq m<M}$
  - deux bases : $(\psi_m)_{0\leq m<M}$ and $(\psi_m^H)_{0\leq m<M}$
- Conditions d'orthogonalité

$$\sum_{p=0}^{M-1}\{\widetilde{H}_m||\widetilde{G}_m\}(\omega+p\frac{2\pi}{M})\{H^*_{m'}||G^*_{m'}\}(\omega+p\frac{2\pi}{M})=M\delta_{m-m'}.$$

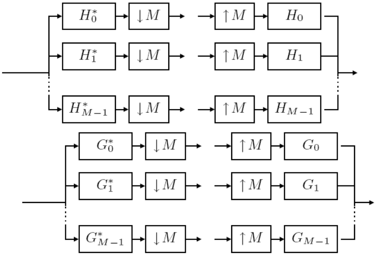
- Équations d'échelle

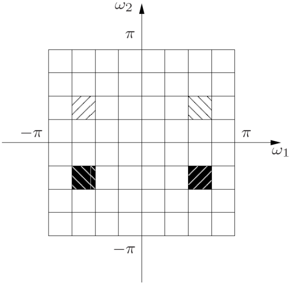
$$\forall m\in\mathbb{N}_M,\qquad \sqrt{M}\widehat{\psi}_m(M\omega)=H_m(\omega)\widehat{\psi}_0(\omega),$$
$$\sqrt{M}\widehat{\psi}^H_m(M\omega)=G_m(\omega)\widehat{\psi}^H_0(\omega),.$$

- Conditions de paires de Hilbert

$$\forall m\in\mathbb{N}^\star_M,\quad \widehat{\psi}^{\mathrm{H}}_m(\omega)=-\imath\,\mathrm{sign}(\omega)\widehat{\psi}_m(\omega)$$

Applications : M-bandes en arbre dual





Ondelettes dual-tree M-bandes Quadrant fréquentiel, 4-bandes

Applications : M-bandes en arbre dual

- Avantages en M-bandes ($M > 2$) :
  - décomposition plus précise dans le domaine fréquentiel
  - plus grande liberté dans le choix des filtres M-bandes $\Longleftrightarrow$ ondelettes symétriques et orthogonales, à support compact

Applications : M-bandes en arbre dual

- Avantages en M-bandes ($M > 2$) :
  - décomposition plus précise dans le domaine fréquentiel
  - plus grande liberté dans le choix des filtres M-bandes $\Longleftrightarrow$ ondelettes symétriques et orthogonales, à support compact
- Avantages de la construction duale :
  - quasi invariance par translation
  - directionnalité en 2D

➤ $M = 2$: [Kingsbury, 1998] [Selesnick, 2001]

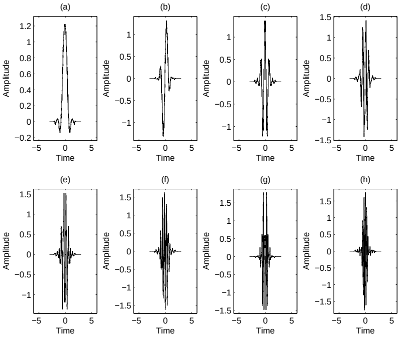
Applications : M-bandes en arbre dual

- Avantages en M-bandes ($M > 2$) :
  - décomposition plus précise dans le domaine fréquentiel
  - plus grande liberté dans le choix des filtres M-bandes $\Longleftrightarrow$ ondelettes symétriques et orthogonales, à support compact
- Avantages de la construction duale :
  - quasi invariance par translation
  - directionnalité en 2D

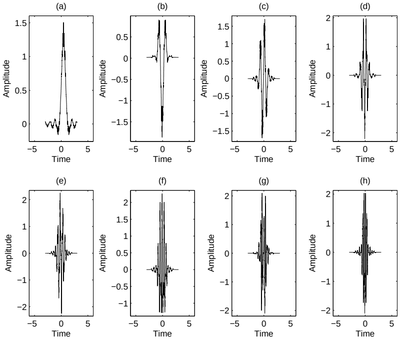
➤ $M = 2$: [Kingsbury, 1998] [Selesnick, 2001]

Transformée géométrique basée sur des bancs de filtres M-bandes ayant une faible redondance (facteur 2) [Chaux *et al.*, 2006]

Bases — 1-D



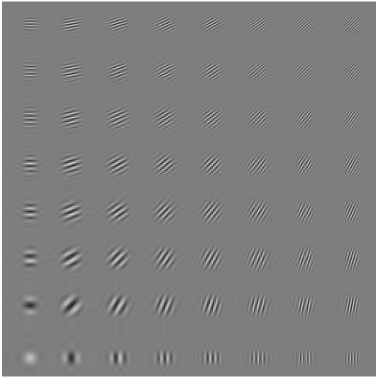
Primal



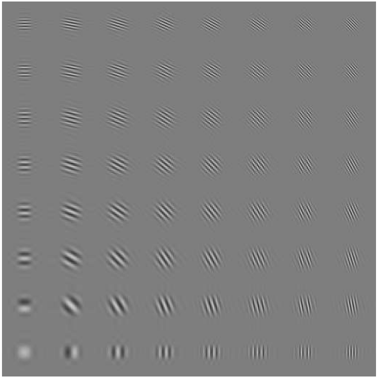
Dual

Figure : En 8 bandes

Bases — 2-D



Dir. positive



Dir. négative

Figure : En 8 bandes

Propriétés du bruit

En présence d'un bruit blanc : pour tout $\boldsymbol{\ell} = (\ell_1, \ell_2) \in \mathbb{Z}^2$,

- Dans le cas $1D$: pour tout $(m, m') \in \{0, ..., M-1\}^2$ et $\ell \in \mathbb{Z}$,

$$\Gamma_{n_{j,m}, n_{j,m'}}(\ell) = \Gamma_{n^{\mathrm{H}}_{j,m}, n^{\mathrm{H}}_{j,m'}}(\ell) = \sigma^2 \delta_{m-m'} \delta_\ell$$

$$\Gamma_{n_{j,m}, n^{\mathrm{H}}_{j,m'}}(\ell) = \sigma^2 \gamma_{\psi_m, \psi^{\mathrm{H}}_{m'}}(-\ell)\,,$$

Propriétés du bruit

En présence d'un bruit blanc : pour tout $\boldsymbol{\ell} = (\ell_1, \ell_2) \in \mathbb{Z}^2$,

- Dans le cas $1D$: pour tout $(m, m') \in \{0, \dots, M-1\}^2$ et $\ell \in \mathbb{Z}$,

$$\Gamma_{n_{j,m}, n_{j,m'}}(\ell) = \Gamma_{n^{\mathrm{H}}_{j,m}, n^{\mathrm{H}}_{j,m'}}(\ell) = \sigma^2 \delta_{m-m'} \delta_\ell$$

$$\Gamma_{n_{j,m}, n^{\mathrm{H}}_{j,m'}}(\ell) = \sigma^2 \gamma_{\psi_m, \psi^{\mathrm{H}}_{m'}}(-\ell),$$

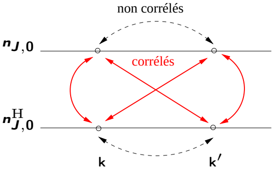
- Dans le cas $2D$: pour tout $(\mathbf{m}, \mathbf{m}') \in \{0, \dots, M-1\}^2 \times \{0, \dots, M-1\}^2$ et $\boldsymbol{\ell} \in \mathbb{Z}^2$,

$$\Gamma_{n_{j,\mathbf{m}}, n_{j,\mathbf{m}'}}(\boldsymbol{\ell}) = \Gamma_{n^{\mathrm{H}}_{j,\mathbf{m}}, n^{\mathrm{H}}_{j,\mathbf{m}'}}(\boldsymbol{\ell}) = \sigma^2 \delta_{m_1 - m'_1} \delta_{m_2 - m'_2} \delta_{\ell_1} \delta_{\ell_2}$$

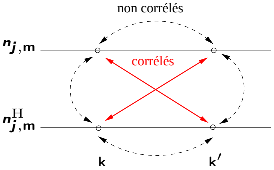
$$\Gamma_{n_{j,\mathbf{m}}, n^{\mathrm{H}}_{j,\mathbf{m}'}}(\boldsymbol{\ell}) = \sigma^2 \gamma_{\psi_{m_1}, \psi^{\mathrm{H}}_{m'_1}}(-\ell_1) \gamma_{\psi_{m_2}, \psi^{\mathrm{H}}_{m'_2}}(-\ell_2)$$

Rque : La comb. lin. des sous-bandes annule les corrélations inter-arbres à une position spatiale **k** donnée mais crée des corrélations spatiales.

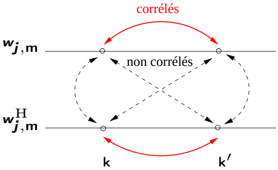
Résumé



Coefficients d'approximation



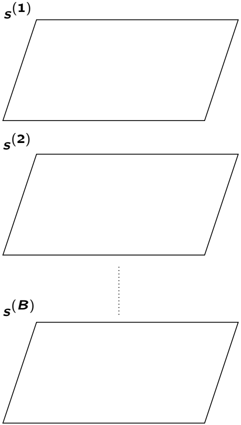
Coefficients de détails ($\mathbf{m} \neq \mathbf{0}$)

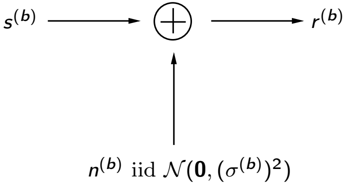


Coefficients de détails avec $m_1 m_2 \neq 0$

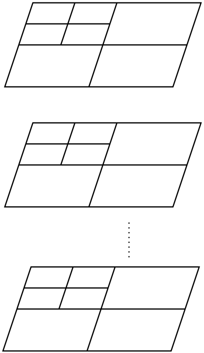
Schéma des corrélations inter-arbres pour une sous-bande $(j, \mathbf{m})$ donnée, avant (haut) et après (bas) isométrie [Chaux, 2007].

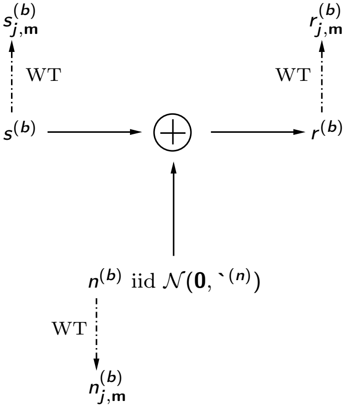
Imagerie multi-canaux





Imagerie multi-canaux

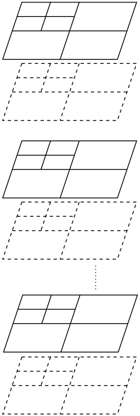


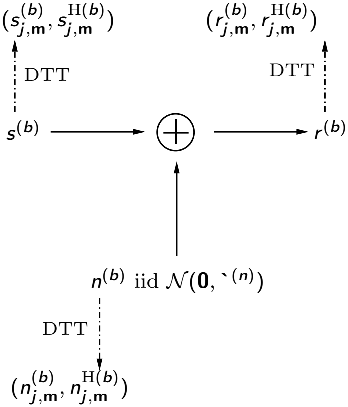


Objectif :

Construire un estimateur $\hat{s}^{(b)}_{j,\mathbf{m}}$ des coef. d'ondelettes $s^{(b)}_{j,\mathbf{m}}$ de $s^{(b)}$ à partir des coefficients d'ondelettes observés $r^{(b')}_{j,\mathbf{m}}$ de tous les canaux b'.

Imagerie multi-canaux



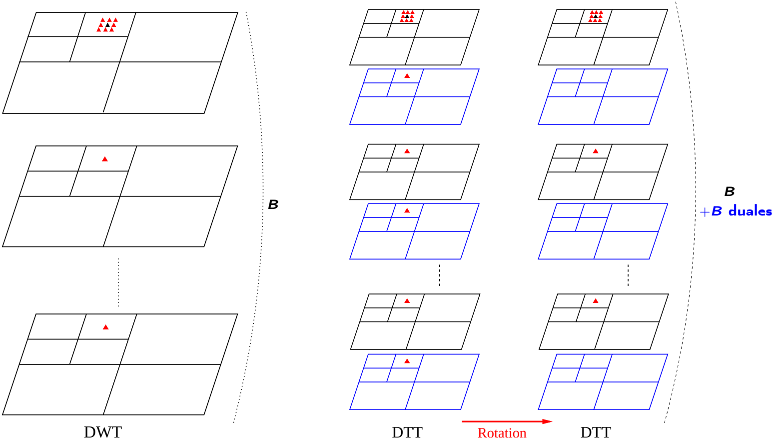


Objectif :

Construire un estimateur $\hat{s}_{j,\mathbf{m}}^{(b)}$ (resp. $\hat{s}_{j,\mathbf{m}}^{\mathrm{H}(b)}$) des coef. issus de la DTT $s_{j,\mathbf{m}}^{(b)}$ (resp. $s_{j,\mathbf{m}}^{\mathrm{H}(b)}$) de $s^{(b)}$ (resp. $s^{\mathrm{H}(b)}$) à partir des coef. observés $r_{j,\mathbf{m}}^{(b')}$ et $r_{j,\mathbf{m}}^{\mathrm{H}(b')}$.

Voisinage

$\bar{\mathbf{r}}_{j,\mathbf{m}}^{(b)}(\mathbf{k})$: coefficients observés utilisés au débruitage de $s_{j,\mathbf{m}}^{(b)}(\mathbf{k})$ (ROV). On a choisi un voisinage de d coefficients combinant des informations *inter-composantes et spatiales* :



Estimateurs: point de départ

- Neighblock [Cai et Silverman, 2001]:
  Étude asymtotique minimax

$$\overset{\curlywedge}{\mathbf{s}}^{(b)}_{j,\mathbf{m}}(\mathbf{k}) = \left(\frac{\|\mathbf{r}^{(b)}_{j,\mathbf{m}}(\mathbf{k})\|^{2} - \lambda}{\|\mathbf{r}^{(b)}_{j,\mathbf{m}}(\mathbf{k})\|^{2}}\right)_{+} \mathbf{r}^{(b)}_{j,\mathbf{m}}(\mathbf{k})$$

Estimateurs: point de départ

- Neighblock [Cai et Silverman, 2001]:
  Étude asymtotique minimax

$$\overset{\curlywedge}{\mathbf{s}}^{(b)}_{j,\mathbf{m}}(\mathbf{k}) = \left(\frac{\|\mathbf{r}^{(b)}_{j,\mathbf{m}}(\mathbf{k})\|^2 - \lambda}{\|\mathbf{r}^{(b)}_{j,\mathbf{m}}(\mathbf{k})\|^2}\right)_+ \mathbf{r}^{(b)}_{j,\mathbf{m}}(\mathbf{k})$$

- Seuillage bivarié [Şendur et Selesnick, 2002]:
  Approche bayésienne

$$\overset{\curlywedge}{s}^{(b)}_{j,\mathbf{m}}(\mathbf{k}) = \left(\frac{\sqrt{\left(r^{(b)}_{j,\mathbf{m}}(\mathbf{k})\right)^2 + \left(r^{(b)}_{j+1,\mathbf{m}}(\lceil \frac{\mathbf{k}}{2} \rceil)\right)^2} - \lambda}{\sqrt{\left(r^{(b)}_{j,\mathbf{m}}(\mathbf{k})\right)^2 + \left(r^{(b)}_{j+1,\mathbf{m}}(\lceil \frac{\mathbf{k}}{2} \rceil)\right)^2}}\right)_+ r^{(b)}_{j,\mathbf{m}}(\mathbf{k})$$

où

$$(f)_+ = \begin{cases} f & \text{si } f > 0 \\ 0 & \text{sinon} \end{cases}$$

$\longrightarrow$ *On omettra les indices $j, \mathbf{m}$ dans la suite.*

Estimateur proposé

- Fonction de seuillage généralisée $\eta_\lambda(\|\mathbf{r}^{(b)}(\mathbf{k})\|^\beta)$, où:

$$\forall \tau \in \mathbb{R}_+, \qquad \boxed{\eta_\lambda(\tau) = \left(\frac{\tau - \lambda}{\tau}\right)_+}$$

et $\beta > 0$ et $\lambda \geq 0$.

Estimateur proposé

- Fonction de seuillage généralisée $\eta_\lambda(\|\bar{\mathbf{r}}^{(b)}(\mathbf{k})\|^\beta)$, où:

$$\forall \tau \in \mathbb{R}_+, \qquad \boxed{\eta_\lambda(\tau) = \left(\frac{\tau - \lambda}{\tau}\right)_+}$$

  et $\beta > 0$ et $\lambda \geq 0$.

-

$$\boxed{\overset{\curlywedge}{s}^{(b)}(\mathbf{k}) = f(\bar{\mathbf{r}}(\mathbf{k})) = \eta_{\lambda^{(b)}}\left(\|\bar{\mathbf{r}}^{(b)}(\mathbf{k})\|^{\beta^{(b)}}\right) \ (\mathbf{q}^{(b)})^\top \bar{\mathbf{r}}^{(b)}(\mathbf{k})}$$

  où $\mathbf{q}^{(b)}$: vecteur de même taille que le ROV, $\beta^{(b)} > 0$, $\lambda^{(b)} \geq 0$.

Calcul des paramètres

- $\lambda^{(b)} \Longleftrightarrow$ valeur de seuil (seuillage doux de $\|\mathbf{r}^{(b)}(\mathbf{k})\|^{\beta^{(b)}}$)

Calcul des paramètres

- $\lambda^{(b)} \Longleftrightarrow$ valeur de seuil (seuillage doux de $\|\mathbf{r}^{(b)}(\mathbf{k})\|^{\beta^{(b)}}$)
- Objectif :
  - Trouver la valeur du seuil $\lambda^{(b)}$
  - Trouver le paramètre exposant $\beta^{(b)}$
  - Trouver le vecteur $\mathbf{q}^{(b)}$ servant à la combinaison linéaire

Calcul des paramètres

- $\lambda^{(b)} \Longleftrightarrow$ valeur de seuil (seuillage doux de $\|\bar{\mathbf{r}}^{(b)}(\mathbf{k})\|^{\beta^{(b)}}$)
- Objectif :
  - Trouver la valeur du seuil $\lambda^{(b)}$
  - Trouver le paramètre exposant $\beta^{(b)}$
  - Trouver le vecteur $\mathbf{q}^{(b)}$ servant à la combinaison linéaire

minimisant l'erreur quadratique moyenne:

$$R(\lambda, \beta, \mathbf{q}) = \mathsf{E}[|s^{(b)}(\mathbf{k}) - \hat{s}^{(b)}(\mathbf{k})|^2]$$
$$= \mathsf{E}[|s^{(b)}(\mathbf{k})|^2] + \mathsf{E}[|f(\bar{\mathbf{r}}^{(b)}(\mathbf{k}))|^2] - 2\mathsf{E}[f(\bar{\mathbf{r}}^{(b)}(\mathbf{k}))s^{(b)}(\mathbf{k})].$$

Calcul des paramètres

- $\lambda^{(b)} \Longleftrightarrow$ valeur de seuil (seuillage doux de $\|\mathbf{r}^{(b)}(\mathbf{k})\|^{\beta^{(b)}}$)
- Objectif :
  - Trouver la valeur du seuil $\lambda^{(b)}$
  - Trouver le paramètre exposant $\beta^{(b)}$
  - Trouver le vecteur $\mathbf{q}^{(b)}$ servant à la combinaison linéaire

minimisant l'erreur quadratique moyenne:

$$R(\lambda, \beta, \mathbf{q}) = \mathsf{E}[|s^{(b)}(\mathbf{k}) - \overset{\curlywedge}{s}^{(b)}(\mathbf{k})|^2]$$
$$= \mathsf{E}[|s^{(b)}(\mathbf{k})|^2] + \mathsf{E}[|f(\bar{\mathbf{r}}^{(b)}(\mathbf{k}))|^2] - 2\mathsf{E}[f(\bar{\mathbf{r}}^{(b)}(\mathbf{k}))s^{(b)}(\mathbf{k})].$$

❢ $\mathbf{s}^{(b)}$ est inconnu : utilisation du Principe de Stein.

Principe de Stein

$$\forall \mathbf{k} \in \mathbb{Z}^2, \qquad \overline{\mathbf{r}}^{(b)}(\mathbf{k}) = \overline{\mathbf{s}}^{(b)}(\mathbf{k}) + \overline{\mathbf{n}}^{(b)}(\mathbf{k}),$$

$$\overline{\mathbf{r}}^{(b)}(\mathbf{k}) = \begin{bmatrix} r^{(b)}(\mathbf{k}) \\ \tilde{\mathbf{r}}^{(b)}(\mathbf{k}) \end{bmatrix}, \quad \overline{\mathbf{s}}^{(b)}(\mathbf{k}) = \begin{bmatrix} s^{(b)}(\mathbf{k}) \\ \tilde{\mathbf{s}}^{(b)}(\mathbf{k}) \end{bmatrix}, \quad \overline{\mathbf{n}}^{(b)}(\mathbf{k}) = \begin{bmatrix} n^{(b)}(\mathbf{k}) \\ \tilde{\mathbf{n}}^{(b)}(\mathbf{k}) \end{bmatrix}$$

Principe de Stein

$$\forall \mathbf{k} \in \mathbb{Z}^2, \qquad \bar{\mathbf{r}}^{(b)}(\mathbf{k}) = \bar{\mathbf{s}}^{(b)}(\mathbf{k}) + \bar{\mathbf{n}}^{(b)}(\mathbf{k}),$$

$$\bar{\mathbf{r}}^{(b)}(\mathbf{k}) = \begin{bmatrix} r^{(b)}(\mathbf{k}) \\ \tilde{\mathbf{r}}^{(b)}(\mathbf{k}) \end{bmatrix}, \quad \bar{\mathbf{s}}^{(b)}(\mathbf{k}) = \begin{bmatrix} s^{(b)}(\mathbf{k}) \\ \tilde{\mathbf{s}}^{(b)}(\mathbf{k}) \end{bmatrix}, \quad \bar{\mathbf{n}}^{(b)}(\mathbf{k}) = \begin{bmatrix} n^{(b)}(\mathbf{k}) \\ \tilde{\mathbf{n}}^{(b)}(\mathbf{k}) \end{bmatrix}$$

Principe de Stein

Soit $f : \mathbb{R}^d \to \mathbb{R}$ une fonction continue, différentiable presque partout telle que :

$$\forall \boldsymbol{\theta} \in \mathbb{R}^d, \lim_{\|\mathbf{t}\| \to +\infty} f(\mathbf{t}) \exp\left(- \frac{(\mathbf{t} - \boldsymbol{\theta})^\top (\boldsymbol{\Gamma}^{(\bar{\mathbf{n}}^{(b)})})^{-1} (\mathbf{t} - \boldsymbol{\theta})}{2} \right) = 0$$

$$\mathsf{E}[|f(\bar{\mathbf{r}}^{(b)}(\mathbf{k}))|^2] < +\infty \quad \text{et} \quad \mathsf{E}\left[\left\| \frac{\partial f(\bar{\mathbf{r}}^{(b)}(\mathbf{k}))}{\partial \bar{\mathbf{r}}^{(b)}(\mathbf{k})} \right\|\right] < +\infty.$$

On a, $\mathsf{E}[f(\bar{\mathbf{r}}^{(b)}(\mathbf{k})) s^{(b)}(\mathbf{k})] =$

$$\mathsf{E}[f(\bar{\mathbf{r}}^{(b)}(\mathbf{k})) r^{(b)}(\mathbf{k})] - \mathsf{E}\left[\frac{\partial f(\bar{\mathbf{r}}^{(b)}(\mathbf{k}))}{\partial \bar{\mathbf{r}}^{(b)}(\mathbf{k})} \right]^\top \mathsf{E}[\bar{\mathbf{n}}^{(b)} n^{(b)}].$$

Calcul des paramètres

- Dans notre cas $f \,:\, \mathbf{u} \mapsto \eta_{\lambda^{(b)}}(\|\mathbf{u}\|^{\beta^{(b)}})\,(\mathbf{q}^{(b)})^\top \mathbf{u}$

Calcul des paramètres

- Dans notre cas $f\ :\ \mathbf{u} \mapsto \eta_{\lambda^{(b)}}(\|\mathbf{u}\|^{\beta^{(b)}})\ (\mathbf{q}^{(b)})^\top \mathbf{u}$
- Afin de limiter la complexité de calcul, $\mathbf{q}^{(b)}$ peut être contraint à un sous-espace de dimension $d' \leq d$. Soit $\mathbf{P}^{(b)} \in \mathbb{R}^{d \times d'}$ une matrice dont les colonnes forment une base de ce sous-espace ; $\mathbf{q}^{(b)} = \mathbf{P}^{(b)}\mathbf{a}^{(b)}$ où $\mathbf{a}^{(b)} \in \mathbb{R}^{d'}$. Exemple :

$$\mathbf{P}^{(b)} = \begin{bmatrix} \mathbf{I}_{d'} \\ \mathbf{O} \end{bmatrix}$$

$\mathbf{I}_{d'}$: matrice identité de taille $d' \times d'$

Calcul des paramètres

- Dans notre cas $f : \mathbf{u} \mapsto \eta_{\lambda^{(b)}}(\|\mathbf{u}\|^{\beta^{(b)}})\ (\mathbf{q}^{(b)})^\top \mathbf{u}$
- Afin de limiter la complexité de calcul, $\mathbf{q}^{(b)}$ peut être contraint à un sous-espace de dimension $d' \leq d$. Soit $\mathbf{P}^{(b)} \in \mathbb{R}^{d \times d'}$ une matrice dont les colonnes forment une base de ce sous-espace ; $\mathbf{q}^{(b)} = \mathbf{P}^{(b)}\mathbf{a}^{(b)}$ où $\mathbf{a}^{(b)} \in \mathbb{R}^{d'}$. Exemple :

$$\mathbf{P}^{(b)} = \begin{bmatrix} \mathbf{I}_{d'} \\ \mathbf{O} \end{bmatrix}$$

  $\mathbf{I}_{d'}$: matrice identité de taille $d' \times d'$
- Dans ce cas $f : \mathbf{u} \mapsto \eta_{\lambda^{(b)}}(\|\mathbf{u}\|^{\beta^{(b)}})\ (\mathbf{a}^{(b)})^\top (\mathbf{P}^{(b)})^\top \mathbf{u}$

Calcul des paramètres

- Dans notre cas $f : \mathbf{u} \mapsto \eta_{\lambda^{(b)}}(\|\mathbf{u}\|^{\beta^{(b)}})\ (\mathbf{q}^{(b)})^\top \mathbf{u}$
- Afin de limiter la complexité de calcul, $\mathbf{q}^{(b)}$ peut être contraint à un sous-espace de dimension $d' \leq d$. Soit $\mathbf{P}^{(b)} \in \mathbb{R}^{d \times d'}$ une matrice dont les colonnes forment une base de ce sous-espace ; $\mathbf{q}^{(b)} = \mathbf{P}^{(b)}\mathbf{a}^{(b)}$ où $\mathbf{a}^{(b)} \in \mathbb{R}^{d'}$. Exemple :

$$\mathbf{P}^{(b)} = \begin{bmatrix} \mathbf{I}_{d'} \\ \mathbf{O} \end{bmatrix}$$

  $\mathbf{I}_{d'}$: matrice identité de taille $d' \times d'$
- Dans ce cas $f : \mathbf{u} \mapsto \eta_{\lambda^{(b)}}(\|\mathbf{u}\|^{\beta^{(b)}})\ (\mathbf{a}^{(b)})^\top (\mathbf{P}^{(b)})^\top \mathbf{u}$
- Les paramètres à déterminer sont alors $\lambda^{(b)}, \beta^{(b)}$ et $\mathbf{a}^{(b)}$ pour $\mathbf{P}^{(b)}$ donnée.

Algorithme

- **Etape 1** Risque quadratique en $\mathbf{a}$: calcul de $\mathbf{a}$ pour λ et β fixés

Algorithme

- **Etape 1** Risque quadratique en $\mathbf{a}$: calcul de $\mathbf{a}$ pour λ et β fixés
- **Etape 2** Risque quadratique par morceaux en $\lambda^{(b)}$
  $R(\lambda, \beta, \mathbf{a}) = \mathrm{E}[\alpha_2(\mathbf{k})(\lambda^{(b)})^2 + \alpha_1(\mathbf{k})\lambda^{(b)} + \alpha_0(\mathbf{k})]$
  Généralisation de SUREShrink [Donoho *et al.*, 1995].
  Réordonnancement : $\|\bar{\mathbf{r}}(\mathbf{k}_1)\| \geq \|\bar{\mathbf{r}}(\mathbf{k}_2)\| \geq \ldots \geq \|\bar{\mathbf{r}}(\mathbf{k}_K)\|$
  $\lambda \in I_{i_0}$ avec $i_0 \in \{1, \ldots, K+1\}$ et

$$I_{i_0} = \begin{cases} \left[\|\bar{\mathbf{r}}(\mathbf{k}_1)\|^\beta, \infty\right) & \text{si } i_0 = 1 \\ \left[\|\bar{\mathbf{r}}(\mathbf{k}_{i_0})\|^\beta, \|\bar{\mathbf{r}}(\mathbf{k}_{i_0-1})\|^\beta\right) & \text{si } i_0 \in \{2, \ldots, K\} \\ \left[0, \|\bar{\mathbf{r}}(\mathbf{k}_K)\|^\beta\right) & \text{si } i_0 = K+1. \end{cases}$$

Pour un $\beta^{(b)}$ fixé on détermine le $\lambda^{(b)}$ optimal

Algorithme

- **Etape 1** Risque quadratique en $\mathbf{a}$: calcul de $\mathbf{a}$ pour λ et β fixés
- **Etape 2** Risque quadratique par morceaux en $\lambda^{(b)}$
  $R(\lambda, \beta, \mathbf{a}) = \mathsf{E}[\alpha_2(\mathbf{k})(\lambda^{(b)})^2 + \alpha_1(\mathbf{k})\lambda^{(b)} + \alpha_0(\mathbf{k})]$
  Généralisation de SUREShrink [Donoho *et al.*, 1995].
  Réordonnancement : $\|\bar{\mathbf{r}}(\mathbf{k}_1)\| \geq \|\bar{\mathbf{r}}(\mathbf{k}_2)\| \geq \ldots \geq \|\bar{\mathbf{r}}(\mathbf{k}_K)\|$
  $\lambda \in I_{i_0}$ avec $i_0 \in \{1, \ldots, K+1\}$ et
  $$I_{i_0} = \begin{cases} \left[\|\bar{\mathbf{r}}(\mathbf{k}_1)\|^\beta, \infty\right) & \text{si } i_0 = 1 \\ \left[\|\bar{\mathbf{r}}(\mathbf{k}_{i_0})\|^\beta, \|\bar{\mathbf{r}}(\mathbf{k}_{i_0-1})\|^\beta\right) & \text{si } i_0 \in \{2, \ldots, K\} \\ \left[0, \|\bar{\mathbf{r}}(\mathbf{k}_K)\|^\beta\right) & \text{si } i_0 = K+1. \end{cases}$$
  Pour un $\beta^{(b)}$ fixé on détermine le $\lambda^{(b)}$ optimal
- **Etape 3** Recherche exhaustive sur un ensemble réduit de valeurs pour trouver le meilleur $\beta^{(b)}$

Comparaisons

Acronym	Description	Acronym	Description
Biv.	Seuillage bivarié [Şendur, Selesnick, 2002]	Méthodes multivariées	
BLS-GSM	Bayesian Least Squares Gaussian Scale Mixture DWT décimée [Portilla, 2003]	ProbShrink $(.\times.)$	Multiv. (images 3-bandes) DWT décimée voisinage $(.\times.)$ [Pizurica, 2006]
BLS-GSM + parent	BLS-GSM DWT décimée + coefficient parent [Portilla, 2003]	ProbShrink red. $(.\times.)$	Multiv. (images 3-bandes) DWT non décimée + voisinage $(.\times.)$ [Pizurica, 2006]
BLS-GSM red.	BLS-GSM + pyramide orientée (transformée redondante) [Portilla, 2003]	Surevect	Appr. SURE étendue DWT décimée [Benazza, 2005]
Curvelets	Curvelets + Estim. par bloc redondance de 7.5 [Candès, 2006]		

Résultats comparatifs

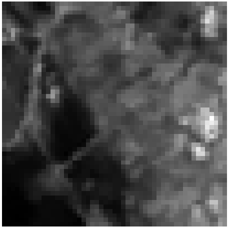
Non redondant

Canal	$\mathrm{SNR}_{\mathrm{init}}$	Biv	ProbShrink (3×3)	BLS-GSM	BLS-GSM + parent	Surevect	Méthode proposée
$b=1$	4.664	11.12	11.18	11.32	11.42	12.95	13.24
$b=2$	5.653	11.60	11.63	11.81	11.90	12.99	13.34
$b=3$	4.926	12.82	12.77	13.01	13.11	13.30	13.65
Moy.	5.081	11.85	11.86	12.05	12.14	13.08	**13.41**
$b=1$	14.66	17.32	17.05	17.67	17.81	19.16	19.71
$b=2$	15.65	18.01	17.65	18.36	18.50	19.35	19.94
$b=3$	14.93	18.32	17.93	18.64	18.74	18.61	19.27
Moy.	15.08	17.88	17.54	18.22	18.35	19.04	**19.64**

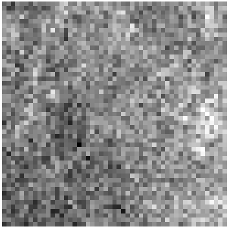
Redondant

Canal	$\mathrm{SNR}_{\mathrm{init}}$	Curvelets	BLS-GSM red + parent	ProbShrink red (3×3)	ProbShrink red (1×1)	Méthode proposée
$b=1$	4.664	11.16	12.21	12.36	12.84	13.57
$b=2$	5.653	11.64	12.69	12.77	13.14	13.65
$b=3$	4.925	12.92	13.88	13.88	14.00	13.91
Moy.	5.081	11.91	12.92	13.00	13.33	**13.71**
$b=1$	14.66	16.87	18.47	18.22	18.95	19.94
$b=2$	15.65	17.50	19.12	18.80	19.38	20.16
$b=3$	14.93	18.31	19.53	19.24	19.28	19.53
Moy.	15.08	17.56	19.04	18.76	19.20	**19.88**

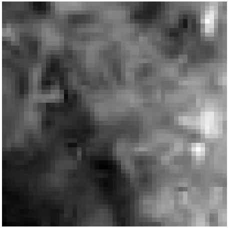
Résultats comparatifs



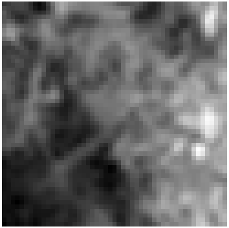
Originale



Bruitée



Probshrink



Dual-tree

Conclusions

- Applications de débruitage

Conclusions

- Applications de débruitage
  - combinaison d'estimateurs statistiques et décompositions directionnelles

Conclusions

- Applications de débruitage
  - combinaison d'estimateurs statistiques et décompositions directionnelles
  - extensions à des bruits de Poisson, des mélanges, au défloutage
- Codes, démos
  - https://www.projet-plume.org/relier/surelet-deconv
  - http://www.greyc.ensicaen.fr/~jfadili/software.html
  - http://www.laurent-duval.eu/misc-research-codes.html
- Perspectives

Conclusions

- Applications de débruitage
  - combinaison d'estimateurs statistiques et décompositions directionnelles
  - extensions à des bruits de Poisson, des mélanges, au défloutage
- Codes, démos
  - `https://www.projet-plume.org/relier/surelet-deconv`
  - `http://www.greyc.ensicaen.fr/~jfadili/software.html`
  - `http://www.laurent-duval.eu/misc-research-codes.html`
- Perspectives
  - erreurs : de vraies mesures pratiques

Conclusions

- Applications de débruitage
  - combinaison d'estimateurs statistiques et décompositions directionnelles
  - extensions à des bruits de Poisson, des mélanges, au défloutage
- Codes, démos
  - https://www.projet-plume.org/relier/surelet-deconv
  - http://www.greyc.ensicaen.fr/~jfadili/software.html
  - http://www.laurent-duval.eu/misc-research-codes.html
- Perspectives
  - erreurs : de vraies mesures pratiques
  - hybribation étendue : représentations, statistiques et optimisation

Conclusions

- Applications de débruitage
  - combinaison d'estimateurs statistiques et décompositions directionnelles
  - extensions à des bruits de Poisson, des mélanges, au défloutage
- Codes, démos
  - `https://www.projet-plume.org/relier/surelet-deconv`
  - `http://www.greyc.ensicaen.fr/~jfadili/software.html`
  - `http://www.laurent-duval.eu/misc-research-codes.html`
- Perspectives
  - erreurs : de vraies mesures pratiques
  - hybribation étendue : représentations, statistiques et optimisation
  - meilleure percolation des ondelettes dans les problèmes concrets

Références

- J.-C. Pesquet and D. Leporini. *A new wavelet estimator for image denoising*, in *IEE Sixth Int. Conf. Im. Proc. Appl.*, volume 1, pages 249–253, Dublin, Ireland, Jul. 14-17 1997.
- C. Chaux, L. Duval et J.-C. Pesquet, *Image Analysis Using a Dual-Tree M-Band Wavelet Transform*, IEEE Trans. on Image Proc., Aug. 2006 (pdf).
- C. Chaux, J.-C. Pesquet et L. Duval, *Noise Covariance Properties in Dual-Tree Wavelet Decompositions*, IEEE Trans. on Inf. Theory, Dec. 2007 (pdf).
- C. Chaux, L. Duval, A. Benazza-Benyahia et J.-C. Pesquet, *A Nonlinear Stein Based Estimator for Multichannel Image Denoising*, IEEE Trans. on Signal Proc., Aug. 2008 (pdf).
- J. Gauthier, L. Duval, and J.-C. Pesquet *Optimization of synthesis oversampled complex filter banks*, IEEE Trans. Signal Process., Oct. 2009 (pdf).
- J.-C. Pesquet, A. Benazza-Benyahia et C. Chaux, *A SURE Approach for Digital Signal/Image Deconvolution Problems*, IEEE Trans. on Signal Proc., Dec. 2009 (pdf),
- L. Jacques, L. Duval, C. Chaux, G. Peyré, *A Panorama on Multiscale Geometric Representations, Intertwining Spatial, Directional and Frequency Selectivity*, Signal Processing, Dec. 2011 (pdf).

Références

- D. L. Donoho and I. M. Johnstone *Adapting to unknown smoothness via wavelet shrinkage*, J. Am. Stat. Assoc., Dec. 1995.
- M. Raphan and E. P. Simoncelli *Empirical Bayes least squares estimation without an explicit prior*, Technical Report TR2007-900, Courant Institute of Mathematical Sciences, NYU, 2007.
- F. Luisier, T. Blu, and M. Unser *A new SURE approach to image denoising: Inter-scale orthonormal wavelet thresholding*, IEEE Trans. Image Process., Mar. 2007.
- T. Blu and F. Luisier, *The SURE-LET approach to image denoising*, IEEE Trans. Image Process., Nov. 2007.
- M. Raphan and E.P. Simoncelli *Optimal denoising in redundant representations*, IEEE Trans. Image Process., Aug. 2008.
- M. Raphan and E. P. Simoncelli *Learning least squares estimators without assumed priors or supervision*, Technical Report TR2009-923, Howard Hughes Medical Institute, Center for Neural Science, New York University, Aug. 2009.
- C. Chesneau, M. J. Fadili, and J.-L. Starck, *Stein Block Thresholding For Image Denoising*, Appl. Comp. Harm. Analysis, Jan. 2010
- G. Peyré, *A Review of Adaptive Image Representations*, J. Sel. Topics Signal Proc., 2011