

Statistiques des images naturelles et modélisation

Cours MVA
Yann Gousseau
Telecom Paris

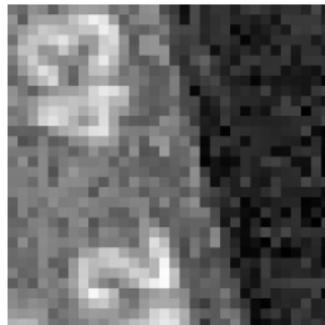
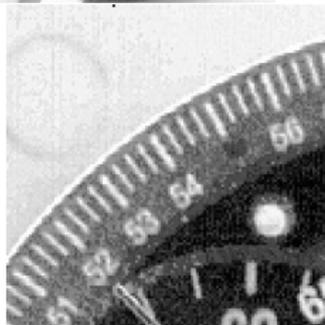
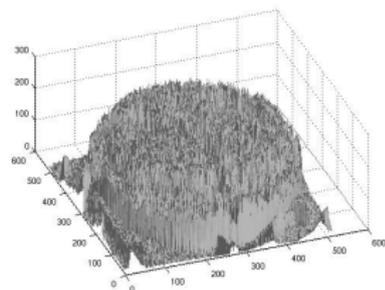


Les signaux bi-dimensionnels correspondant à des images naturelles (images de scènes quelconques) ont-ils des traits récurrents ?

En quoi les images naturelles sont-elles des signaux singuliers ?



Des signaux pour l'oeil



- Nous considérons une image $I : \Omega \subset \mathbb{Z} \rightarrow [0, 1, \dots, G]$ comme un champ aléatoire :
 $\{I(i, j)\}_{i, j \in \Omega}$ est une collection de variables aléatoires.
- Pour une petite image ($256 \times 256 \times 256$) : 2^{524288} réalisations possibles !! (nombre d'atomes dans l'univers estimé à 2^{250}).
- Toutes les réalisations sont-elles équiprobables ?
- Stationnarité (pour I périodisée avec des conditions de symétrie aux frontières de Ω).

- Système sensoriel
- Compression
- Modèles probabilistes (potentiellement utile pour toute approche statistique)
 - Approches bayésiennes de la restauration ou de la détection
 - Tests, méthodes *a contrario*
 - Classification, indexation

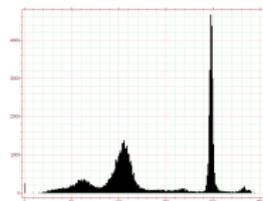
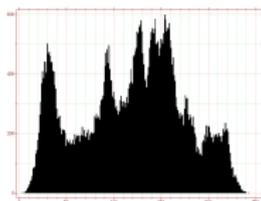
- On cherche à estimer un ensemble $\{a_1, \dots, a_n\}$ sur une image
- Exemples : les a_i sont des labels indiquant la présence d'une structure, les valeurs restaurées des pixels, etc.
- Règle de Bayes :

$$Pr(\{a_i\}|I) = \frac{Pr(I|\{a_i\})Pr(\{a_i\})}{Pr(I)}$$

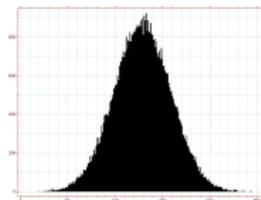
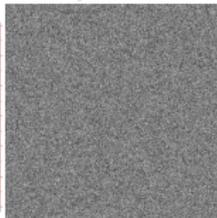
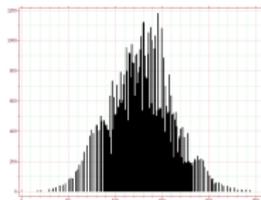
- Si on cherche à maximiser $Pr(\{a_i\}|I)$ (méthode du maximum a posteriori, MAP) il est nécessaire de connaître $Pr(I|\{a_i\})$ et $Pr(\{a_i\})$.

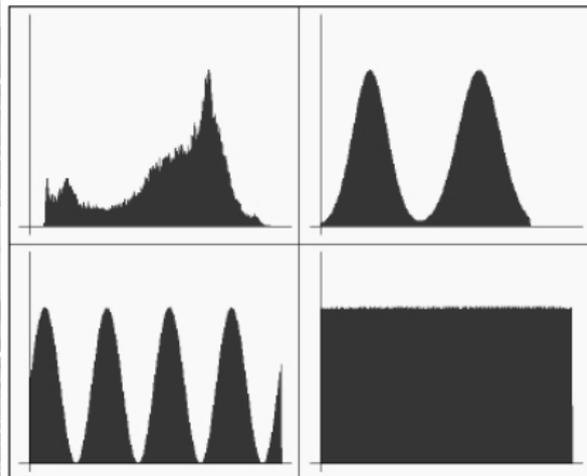
Histogrammes de niveau de gris des images :

-très variable d'une image à l'autre :



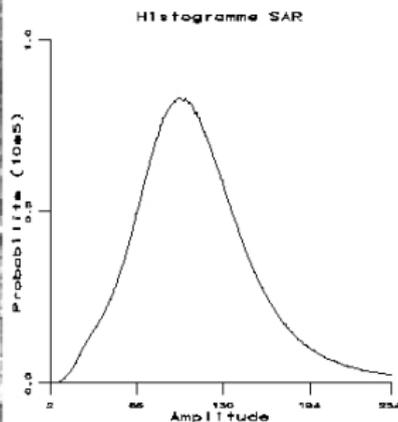
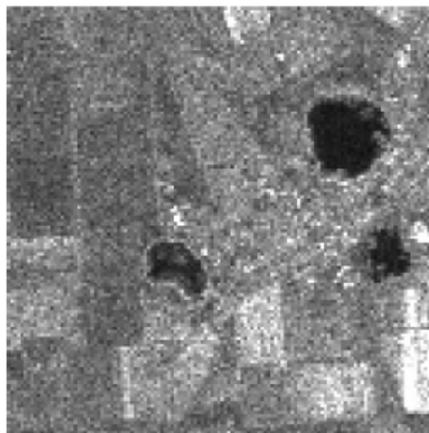
-Peu informatif sur la structure des images :





- Remarque : la situation est très différente pour d'autres types d'images.
- Exemple des images radar : histogramme bien représenté par une loi de Rayleigh

$$p(x) = \sigma^{-2} x \exp\left(-\frac{x^2}{\sigma^2}\right).$$



- Pour une image I , soit \hat{I} sa transformée de Fourier et $P = |\hat{I}|^2$ son spectre.
- Soit $f(t) = \frac{1}{2\pi} \int_0^{2\pi} P(te^{i\theta}) d\theta$
- On observe

$$f(t) \simeq \frac{C}{t^\beta},$$

avec β proche de 2.

- Ce fait est connu depuis les années 50 (signaux de télévision).
- Compatible avec une *invariance par changement d'échelle*.
 $I(\mathbf{x})$ et $I(k\mathbf{x})$ sont distribuées similairement.

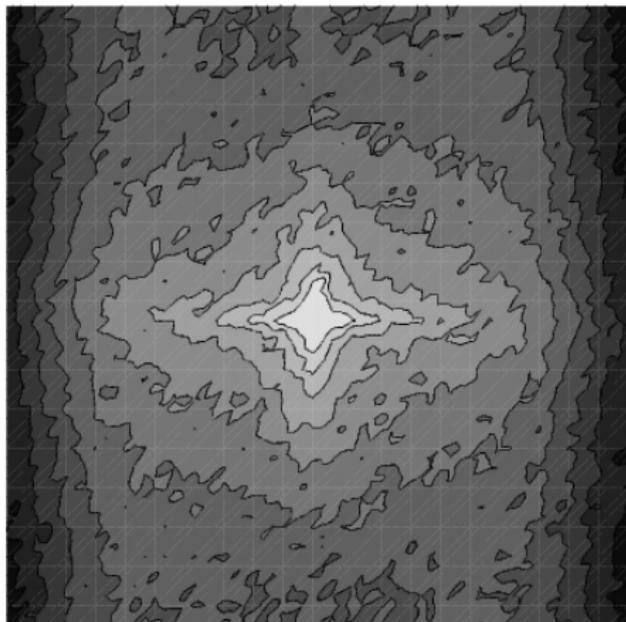


Figure 2: Contour plot of ensemble power spectrum of 45 images taken at focal length of 15mm. Center of figure is $k = 0$. Contours are placed at equal intervals in the logarithm of power, and spatial frequency is plotted on linear axes.

Extrait de [Ruderman 97]

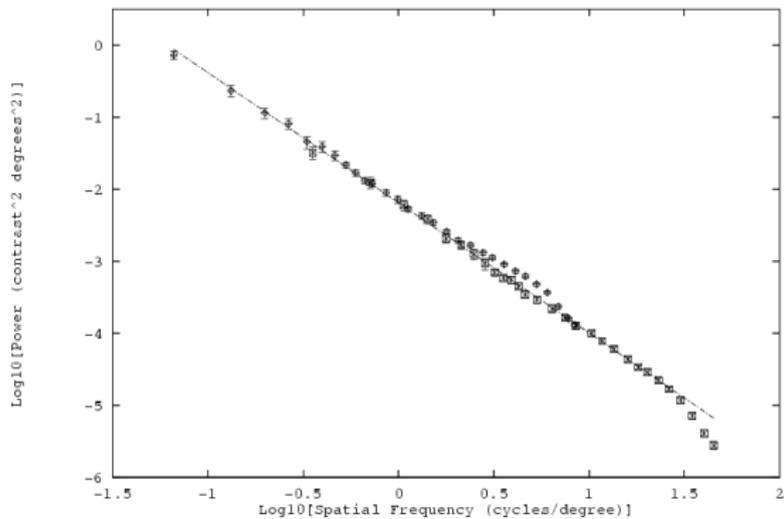
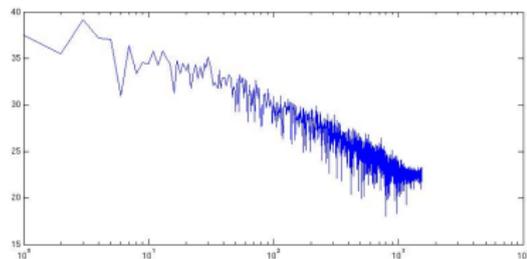
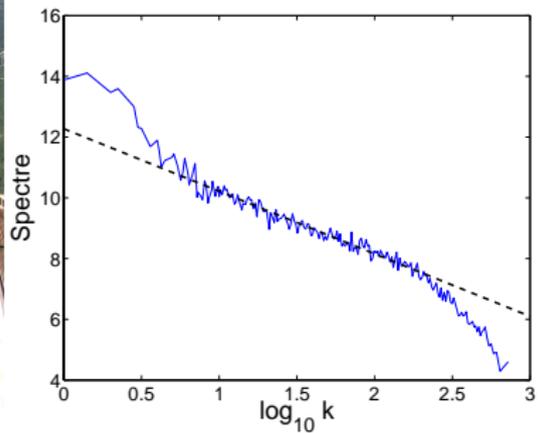
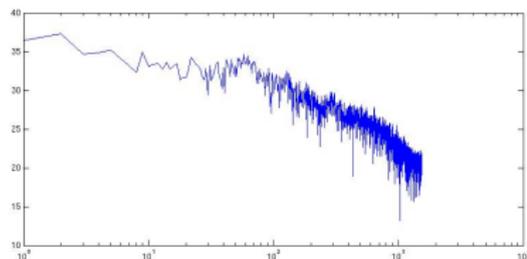


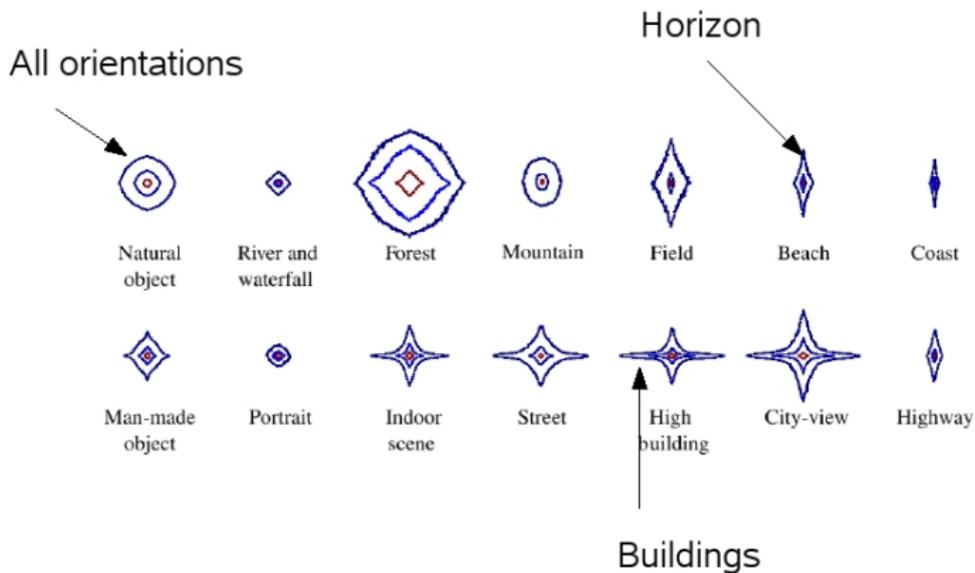
Figure 3: Orientationally averaged power spectrum with standard error bars for 15mm and 80mm focal length data (overlapping), along with regression line fitted as discussed in the text.

Extrait de [Ruderman 97]



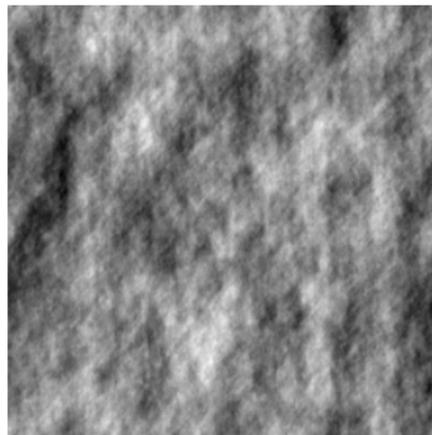






Extrait de [Torralba-Oliva 2003]

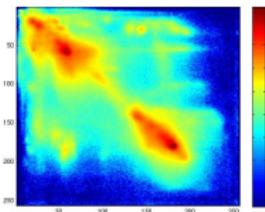
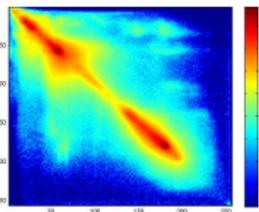
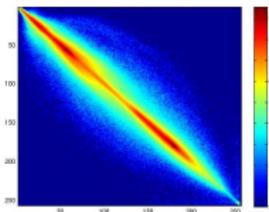
Peu informatif sur la structure des images :



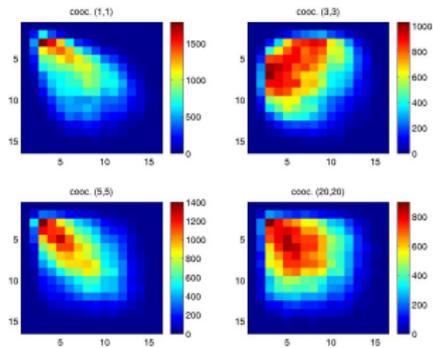
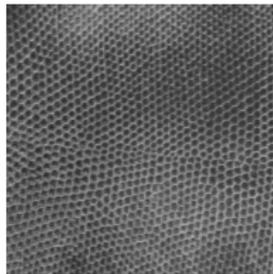
Les images (sauf certaines textures) ne sont pas des champs gaussiens.

- On suppose les images stationnaires
- Covariance : $C(\mathbf{x}) = E(I(0)I(\mathbf{x}))$
- On a $\hat{C} = P$ (formule de Wiener-Khinchin) et (formellement)
 $P(t) = Ct^{-\alpha} \leftrightarrow C(\mathbf{x}) = A + B|\mathbf{x}|^{2-\alpha}$, pour $\alpha \neq 2$,
 $P(t) = Ct^{-2} \leftrightarrow C(\mathbf{x}) = A + B \log(\mathbf{x})$.

- Cooccurrences : $B(\mathbf{x}, i, j) = Pr(I(0) = i \ \& \ I(\mathbf{x}) = j)$.



Cocurrences et textures (Haralick et al., 1973)



- Distributions des sauts $l(\mathbf{x} + \mathbf{y}) - l(\mathbf{x})$ à \mathbf{y} fixé. Proche d'une distribution "gaussienne généralisée" :

$$\rho(x) = C_{\alpha,\eta} \exp(-|\eta x|^\alpha).$$

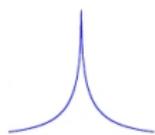


FIG. 1 - $\alpha = 0.5, \eta = 10$



FIG. 2 - $\alpha = 1, \eta = 10$



FIG. 3 - $\alpha = 2, \eta = 10$



FIG. 4 - $\alpha = 4, \eta = 10$

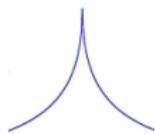


FIG. 5 - $\alpha = 0.5, \eta = 2$



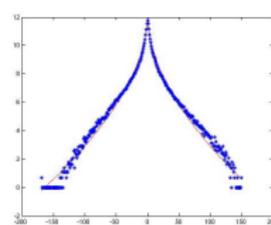
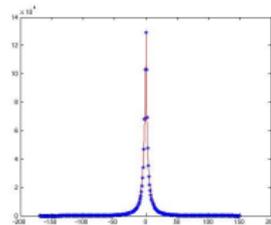
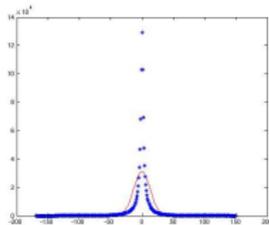
FIG. 6 - $\alpha = 1, \eta = 2$



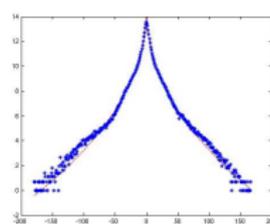
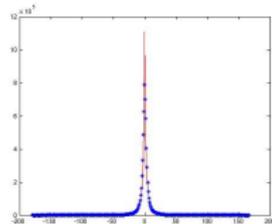
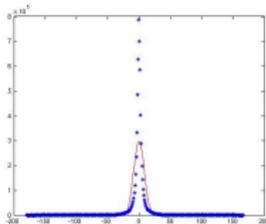
FIG. 7 - $\alpha = 2, \eta = 2$



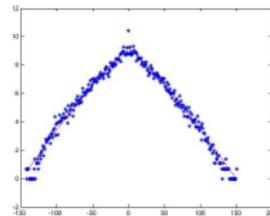
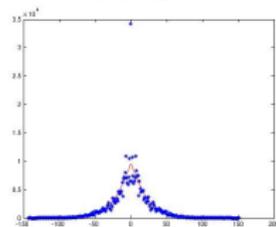
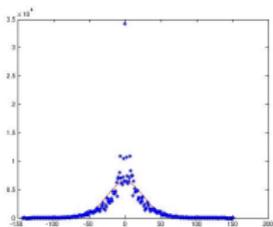
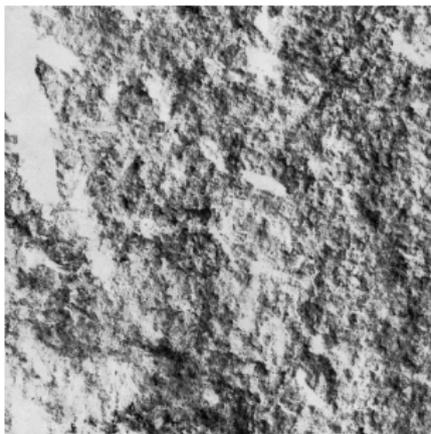
FIG. 8 - $\alpha = 4, \eta = 2$



$\alpha = 0,57$ $\eta = 0,47$ méthode des moments.

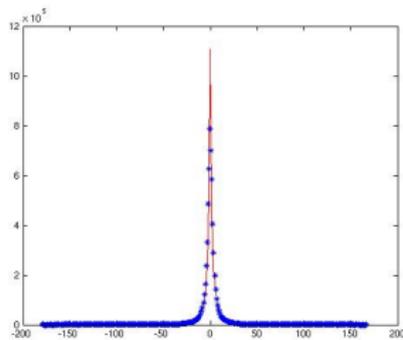


$\alpha = 0,48$ $\eta = 1,59$ méthode des moments.



$\alpha = 1,17$ $\eta = 0,04$ méthode des moments.

- Distributions qualitativement très différentes de distributions gaussiennes
- Intuitivement, ces distributions s'expliquent par la présence de
 - zones homogènes (le fort pic en 0) (effet ciel bleu)
 - bords (les valeurs de module élevé)
- Importance fondamentale du phénomène d'**occultation**.



La non-gaussianité comme mesure de qualité d'une décomposition ?

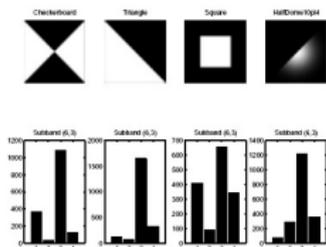


Figure 6. Kurtosis of Fourier(1), wavelet(2), ridgelet(3) and curvelet(4) coefficients; objects having only perfect straight edges, subband $2s = 6$, $j = i = 3$.

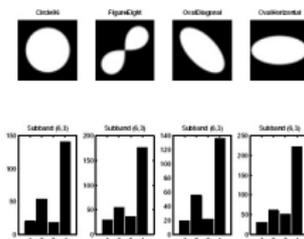


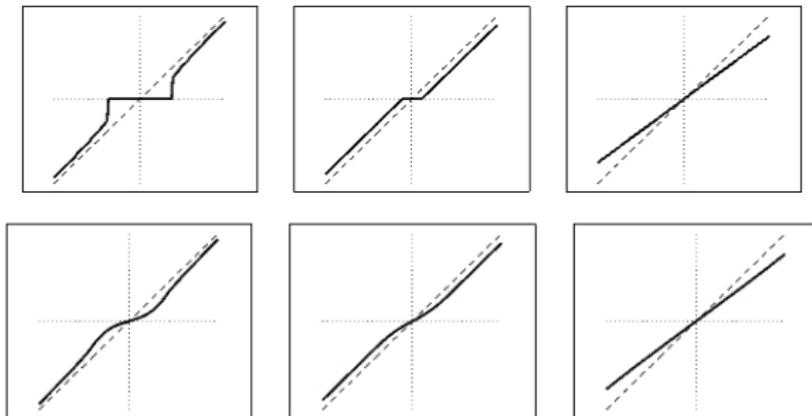
Figure 7. Kurtosis of Fourier (1), wavelet (2), ridgelet (3) and curvelet (4) coefficients; objects with curved edges, subband $2s = 6$, $j = i = 3$.

Extrait de [Donoho-Flesia 2001]

Nombreux travaux sur la décomposition des images en composantes indépendantes ou sur les représentations parcimonieuses des images naturelles (Olshausen-Field 1996, Bell-Sejnowski 1997, Hyvarinen et al 2001, Aharon et al. 2006, Peyré 2009, etc.)

- Les observations faites sur les sauts restent valables pour tout filtre de moyenne nulle ! (Zetsche 1997, Huang 2000)
- En particulier pour les marginales des coefficients d'ondelette (Mallat 1989, Simoncelli)
- Applications
 - débruitage bayésien, compression ([Simoncelli 1997], etc.)
 - détection de contours ([Desolneux et al. 2001], [Heiler-Schnoerr 2005])
 - classification de textures ([Do-Vetterli 2002, Aujol et al. 2003], etc.)
 - restauration variationnelle ([Cho et al. 2012, Tartavel et al. 2016])

Application 1 : débruitage bayésien par ondelettes



Transformation des coefficients d'ondelettes lors d'un débruitage bayésien avec un modèle gaussienne généralisée pour les coefficients.

Haut : max. a posteriori ; bas : moindres carrés bayésiens

De gauche a droite : $\alpha = 0, 5$; $\alpha = 1$; $\alpha = 2$

Variance du bruit = 1/3 de la variance du signal

Extrait de [Simoncelli 2005]

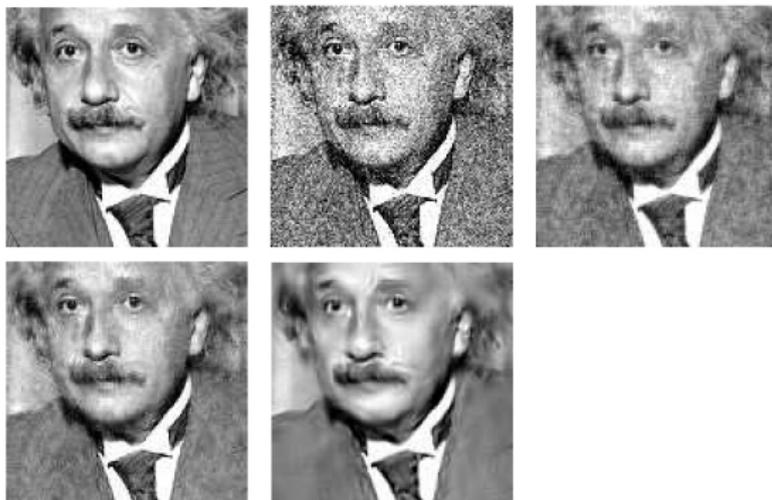


Fig. 10. Denoising examples. **Upper left:** Original 8-bit image (cropped for visibility). **Upper middle:** Contaminated by adding simulated Gaussian white noise, $\sigma = 21.4$, peak-signal-to-noise-ratio (PSNR)=22.06. **Upper right:** Denoised with a Gaussian marginal model (PSNR=27.87). **Lower left:** Denoised with generalized-Gaussian marginal model (PSNR=29.24). **Lower right:** Denoised with a CSM model (PSNR=30.86). All methods were implemented in an overcomplete wavelet domain (see [47, 39]). In each case, the noise variance was assumed known, and the denoising procedure was Bayes least-squares (i.e., the mean of the posterior distribution). Model parameters were estimated by maximizing the likelihood of the data.

Extrait de [Simoncelli 2005]

On s'intéresse au problème de débruitage :

$$y = x + n,$$

avec $n \sim N(0, \sigma_n)$ qui s'écrit dans le domaine ondelettes (base ortho-normée) :

$$\tilde{w} = w + n$$

Cadre bayésien :

$$p(w|\tilde{w}) \propto p(\tilde{w}|w) \cdot p(w)$$

Estimation par MAP (maximum a posteriori)

$$\begin{aligned}\hat{w} &= \arg \max_w p(w|\tilde{w}) \\ &= \arg \min_w [-\log p(\tilde{w}|w) - \log p(w)]\end{aligned}$$

Si on suppose :

- La distribution des coefficients est gaussienne $w \sim N(0, \sigma_w)$ alors

$$\hat{w} = \frac{w}{1 + \frac{\sigma_n^2}{\sigma_w^2}}$$

- La distribution des coefficients est une gaussienne généralisée
→ pas d'expression analytique
approx. numérique : **seuillage**

Mais les réponses à des filtres et coefficients d'ondelettes ne sont pas indépendantes → nécessité de modélisations plus fines.

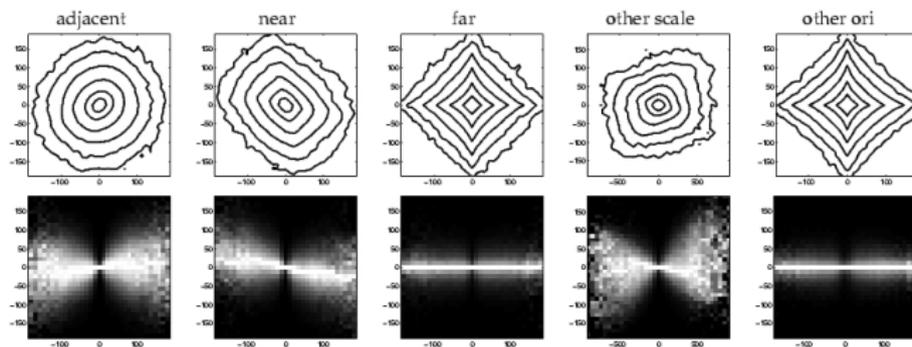


Fig. 8. Empirical joint distributions of wavelet coefficients associated with different pairs of basis functions, for a single image of a New York City street scene (see Fig. 1 for image description). The top row shows joint distributions as contour plots, with lines drawn at equal intervals of log probability. The three leftmost examples correspond to pairs of basis functions at the same scale and orientation, but separated by different spatial offsets. The next corresponds to a pair at adjacent scales (but the same orientation, and nearly the same position), and the rightmost corresponds to a pair at orthogonal orientations (but the same scale and nearly the same position). The bottom row shows corresponding conditional distributions: brightness corresponds to frequency of occurrence, except that each column has been independently rescaled to fill the full range of intensities.

Extrait de [Simoncelli 2005]

On s'intéresse au problème de restauration d'images :

$$y = A.x + n$$

Cadre bayésien :

$$p(x|y) \propto p(y|x).p(x)$$

Estimation par MAP (maximum a posteriori)

$$\begin{aligned}\hat{x} &= \arg \max_x p(x|y) \\ &= \arg \min_x [-\log p(y|x) - \log p(x)] \\ &= \arg \min_x [F(x, y) + R(x)]\end{aligned}$$

avec

F : fidélité

R : a priori \rightarrow régularité

Si on suppose que n est gaussien i.i.d. et que

$$R(x) = \sum_i |\nabla x(i)| := TV(x),$$

alors

$$\hat{x} = \arg \min_x \|A.x - y\|^2 + TV(x),$$

modèle TV-L2 (Rudin-Osher-Fatemi 1992) : voir le cours de restauration.

Le problème :

- Retrouver une image u à partir de $v = \Phi u + n$
- La dégradation Φ est connue (pixels manquants, flou)
- Bruit blanc gaussien n de variance connue

Préservation de la texture :

- Qualité visuelle
- Eviter la dégradation des textures
- Pratique commune : ajout de bruit a posteriori

Texture vs. PSNR :

Lesquelles sont les plus semblables ?

Le problème :

- Retrouver une image u à partir de $v = \Phi u + n$
- La dégradation Φ est connue (pixels manquants, flou)
- Bruit blanc gaussien n de variance connue

Préservation de la texture :

- Qualité visuelle
- Eviter la dégradation des textures
- Pratique commune : ajout de bruit a posteriori

Texture vs. PSNR :

Lesquelles sont les plus semblables ?

Le problème :

- Retrouver une image u à partir de $v = \Phi u + n$
- La dégradation Φ est connue (pixels manquants, flou)
- Bruit blanc gaussien n de variance connue

Préservation de la texture :

- Qualité visuelle
- Eviter la dégradation des textures
- Pratique commune : ajout de bruit a posteriori

Texture vs. PSNR :



Lesquelles sont les plus semblables ?

Modèles variationnels :

Minimisation d'une fonction de coût

$$E_v(u) = \underbrace{R(u)}_{\text{Régularisation}} + \underbrace{F_v(u)}_{\text{Fidélité}}$$

Terme de régularisation :

- Variation totale, $R(u) \propto \sum_x \|\nabla u_x\|_2$ (Rudin-Osher 1992)
- Gradient non local, $R(u) \propto \sum_{x,x'} w_{x,x'} |u_x - u_{x'}|^2$ (Gilboa et al. 2008)
- Parcimonie des patches, $R(u) \propto \min_W \|P_u - DW\|^2 + \|W\|_1$ (Elad et al. 2006)

Terme de fidélité :

- Bruit blanc gaussien, $F_v(u) \propto \|v - \Phi u\|^2$
- Fonction indicatrice $F_v = \iota_{\mathcal{C}} \mathcal{C} = \{u : \|v - \Phi u\|^2 \leq C\sigma^2\}$

Modèles variationnels :

Minimisation d'une fonction de coût

$$E_v(u) = \underbrace{R(u)}_{\text{Régularisation}} + \underbrace{F_v(u)}_{\text{Fidélité}}$$

Terme de régularisation :

- Variation totale, $R(u) \propto \sum_x \|\nabla u_x\|_2$ (Rudin-Osher 1992)
- Gradient non local, $R(u) \propto \sum_{x,x'} w_{x,x'} |u_x - u_{x'}|^2$ (Gilboa et al. 2008)
- Parcimonie des patches, $R(u) \propto \min_W \|P_u - DW\|^2 + \|W\|_1$ (Elad et al. 2006)

Terme de fidélité :

- Bruit blanc gaussien, $F_v(u) \propto \|v - \Phi u\|^2$
- Fonction indicatrice $F_v = \iota_{\mathcal{C}} \mathcal{C} = \{u : \|v - \Phi u\|^2 \leq C\sigma^2\}$

Modèles variationnels :

Minimisation d'une fonction de coût

$$E_v(u) = \underbrace{R(u)}_{\text{Régularisation}} + \underbrace{F_v(u)}_{\text{Fidélité}}$$

Terme de régularisation :

- Variation totale, $R(u) \propto \sum_x \|\nabla u_x\|_2$ (Rudin-Osher 1992)
- Gradient non local, $R(u) \propto \sum_{x,x'} w_{x,x'} |u_x - u_{x'}|^2$ (Gilboa et al. 2008)
- Parcimonie des patches, $R(u) \propto \min_W \|P_u - DW\|^2 + \|W\|_1$ (Elad et al. 2006)

Terme de fidélité :

- Bruit blanc gaussien, $F_v(u) \propto \|v - \Phi u\|^2$
- Fonction indicatrice $F_v = \iota_C \mathcal{C} = \left\{ u : \|v - \Phi u\|^2 \leq C\sigma^2 \right\}$

Formulation :

$$E_v(u) = R(u) + F_v(u) + \tilde{F}_v(u)$$

Avec un terme de fidélité sur le gradient :

$$\tilde{F}_v(u) \propto \mathcal{W}(\mu_{\nabla u}, \mu_0)^2$$

Modèle de gradient :

- ∇u modélisé par une gaussienne généralisée, $d\mu_0(t) \propto e^{-|t/s|^p}$
- Paramètres p et s estimés sur l'image bruitée v

Formulation :

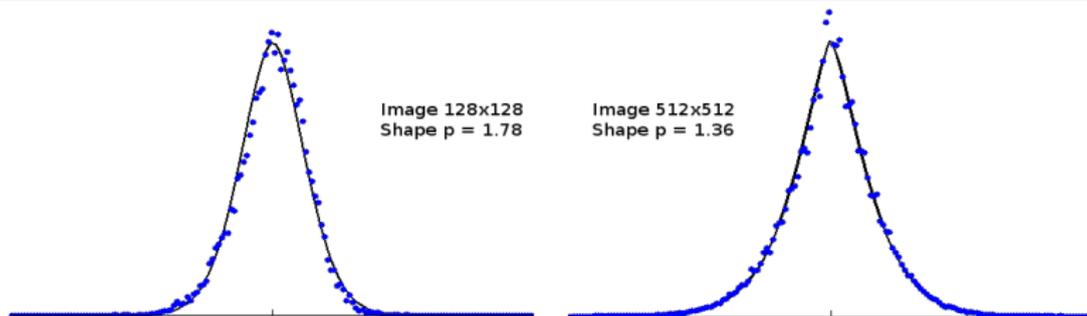
$$E_v(u) = R(u) + F_v(u) + \tilde{F}_v(u)$$

Avec un terme de fidélité sur le gradient :

$$\tilde{F}_v(u) \propto \mathcal{W}(\mu_{\nabla u}, \mu_0)^2$$

Modèle de gradient :

- ∇u modélisé par une gaussienne généralisée, $d\mu_0(t) \propto e^{-|\frac{t}{s}|^p}$
- Paramètres p et s estimés sur l'image bruitée v



Fonction composite :

Fonction de coût $E = F + G + H \circ L$ avec $L = \nabla$,

- $F = \tilde{F}_v$ différentiable et non-convexe
- $G = \iota_C$ proximal ($\text{prox}_G = \text{proj}_C$)
- $H = \|\cdot\|_1$ proximal, avec

$$\text{prox}_{\lambda H}(x) = \arg \min_y \lambda H(y) + \frac{1}{2} \|y - x\|^2$$

Algorithme de minimisation (Condat 2013) :

Algorithme itératif primal-dual :

$$\begin{cases} u_{\ell+1} = \text{prox}_{\tau G}(u_{\ell} - \tau \nabla F(u_{\ell}) - \tau L^* u_{\ell}^*) \\ u_{\ell+1}^* = \text{prox}_{\sigma^* H^*}(u_{\ell}^* + \sigma^* L(2u_{\ell+1} - u_{\ell})) \end{cases}$$

Pas d'assurance de convergence : E est non convexe

Fonction composite :

Fonction de coût $E = F + G + H \circ L$ avec $L = \nabla$,

- $F = \tilde{F}_v$ différentiable et non-convexe
- $G = \iota_C$ proximal ($\text{prox}_G = \text{proj}_C$)
- $H = \|\cdot\|_1$ proximal, avec

$$\text{prox}_{\lambda H}(x) = \arg \min_y \lambda H(y) + \frac{1}{2} \|y - x\|^2$$

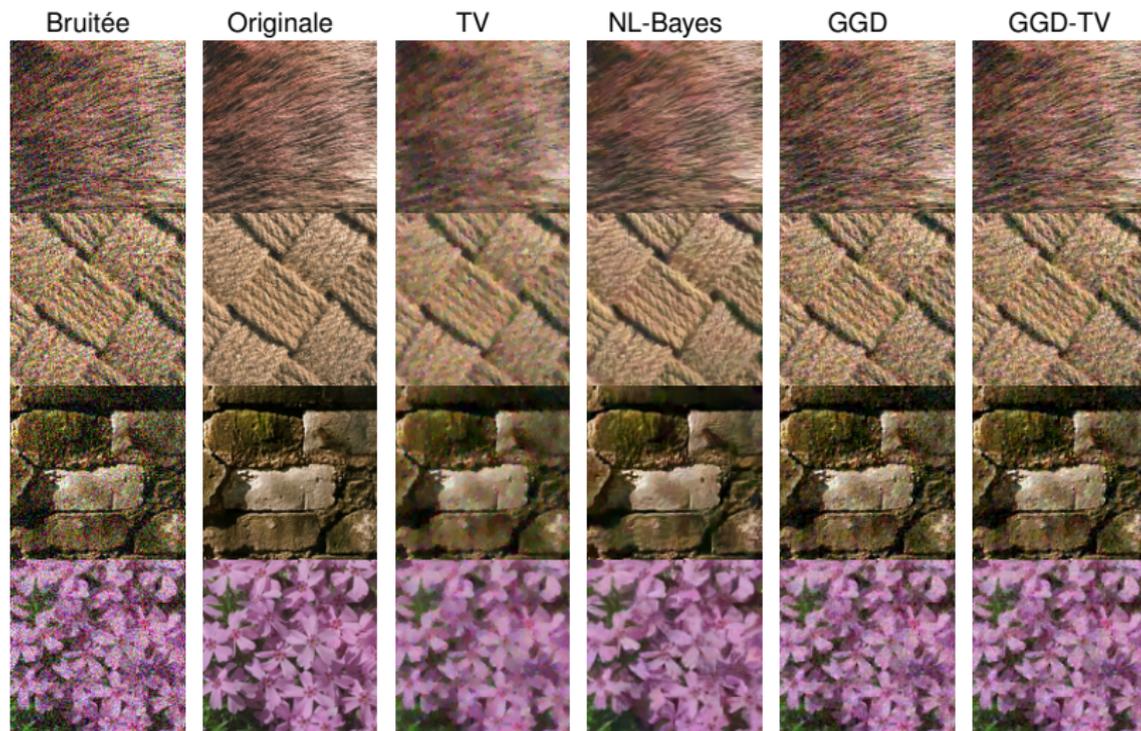
Algorithme de minimisation (Condat 2013) :

Algorithme itératif primal-dual :

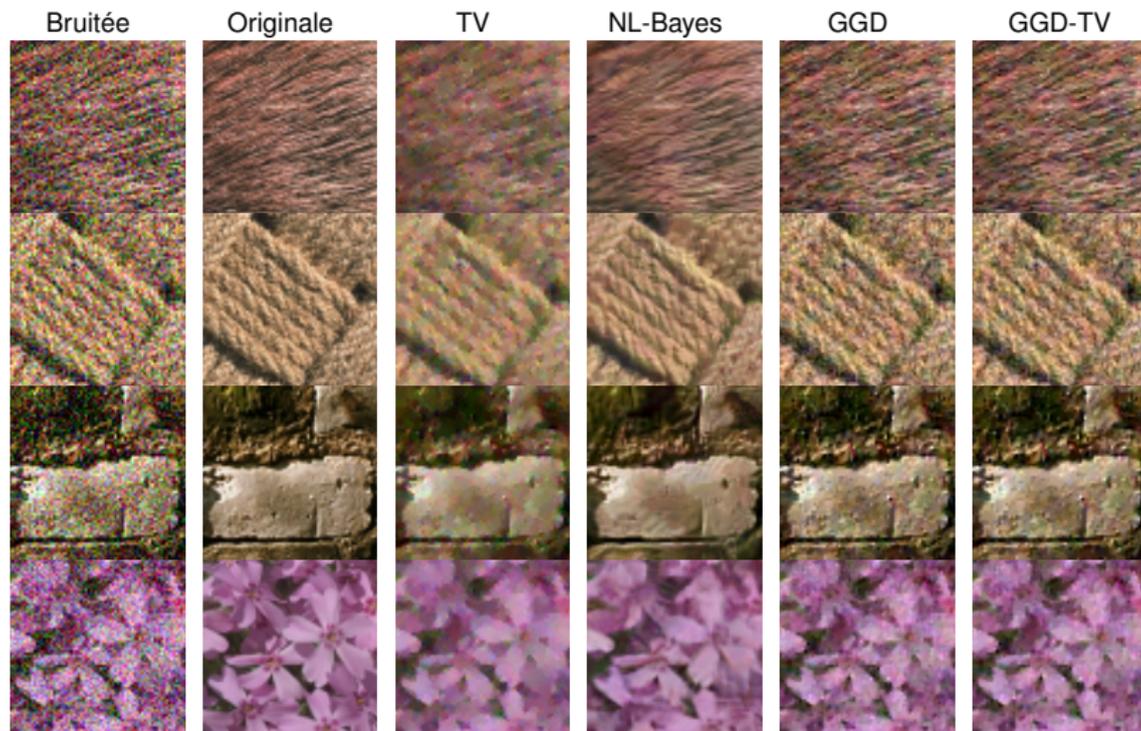
$$\begin{cases} u_{\ell+1} = \text{prox}_{\tau G}(u_{\ell} - \tau \nabla F(u_{\ell}) - \tau L^* u_{\ell}^*) \\ u_{\ell+1}^* = \text{prox}_{\sigma^* H^*}(u_{\ell}^* + \sigma^* L(2u_{\ell+1} - u_{\ell})) \end{cases}$$

Pas d'assurance de convergence : E est non convexe

Résultats de débruitage



Résultats de débruitage (Zoom)

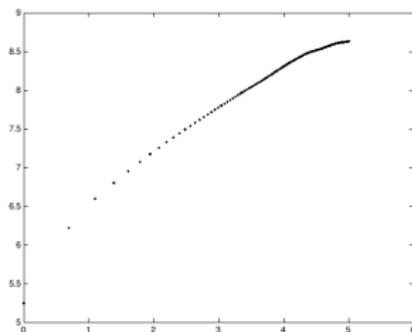


- Pour les réponses $c(y)$ de nombreux filtres à l'échelle y , on observe des comportements du type

$$\text{var}(|c(y)|) \approx Cy^a.$$

Exemples de filtres : sauts, ondelettes, etc.

- Cas des sauts horizontaux à différentes distances, $S(y) = I(\mathbf{x} + (y, 0)) - I(\mathbf{x})$. On observe
 - A nouveau des lois laplaciennes généralisées
 - Une loi d'échelle $\text{var}(S(y)) \approx Cy^a$



- Cohérent avec les formules précédentes pour la covariance (car $\text{var}(I(0) - I(y)) = 2E(I(0))^2 - 2E(I(0)I(y))$).

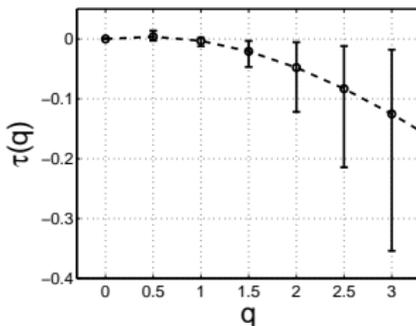
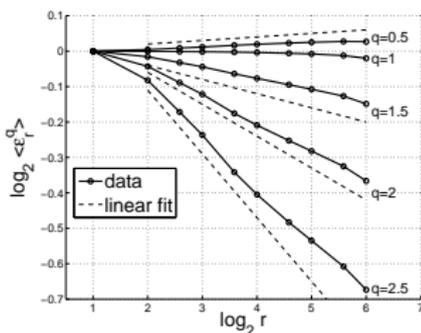
Comportement multi-fractal : dans certains cas on observe aussi (Turiel et al. 1998, Chainais 2007)

$$E(|c(y)|^q) \approx C y^{\xi(q)}$$

avec ξ une fonction non-linéaire.

$$c(y) = \frac{1}{|B(0, y)|} \int_{B(0, y)} |I|$$

$$c(y) = \frac{1}{|B(0, y)|} \int_{B(0, y)} |\nabla I|$$



[Chainais 07]

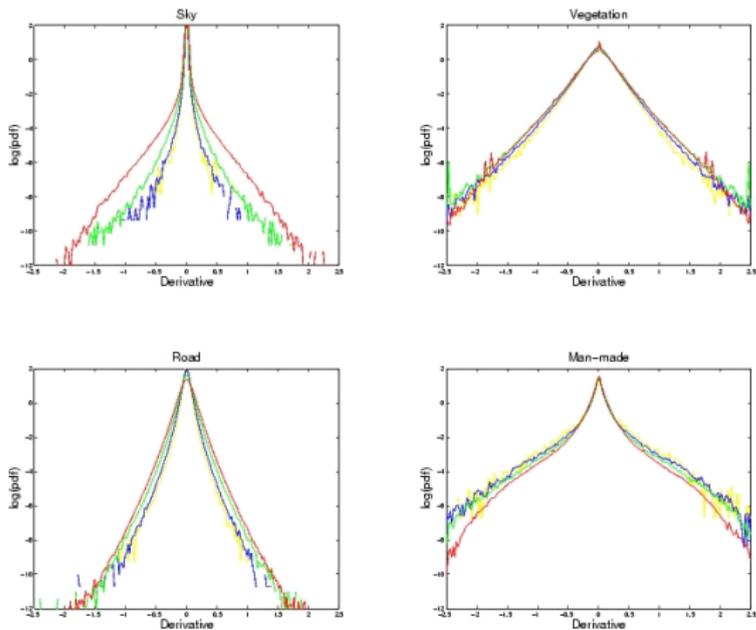


Figure 2.10: $\log(\text{pdf})$ of D at different scales and categories: red = scale 1, green = scale 2, blue = scale 4, yellow = scale 8

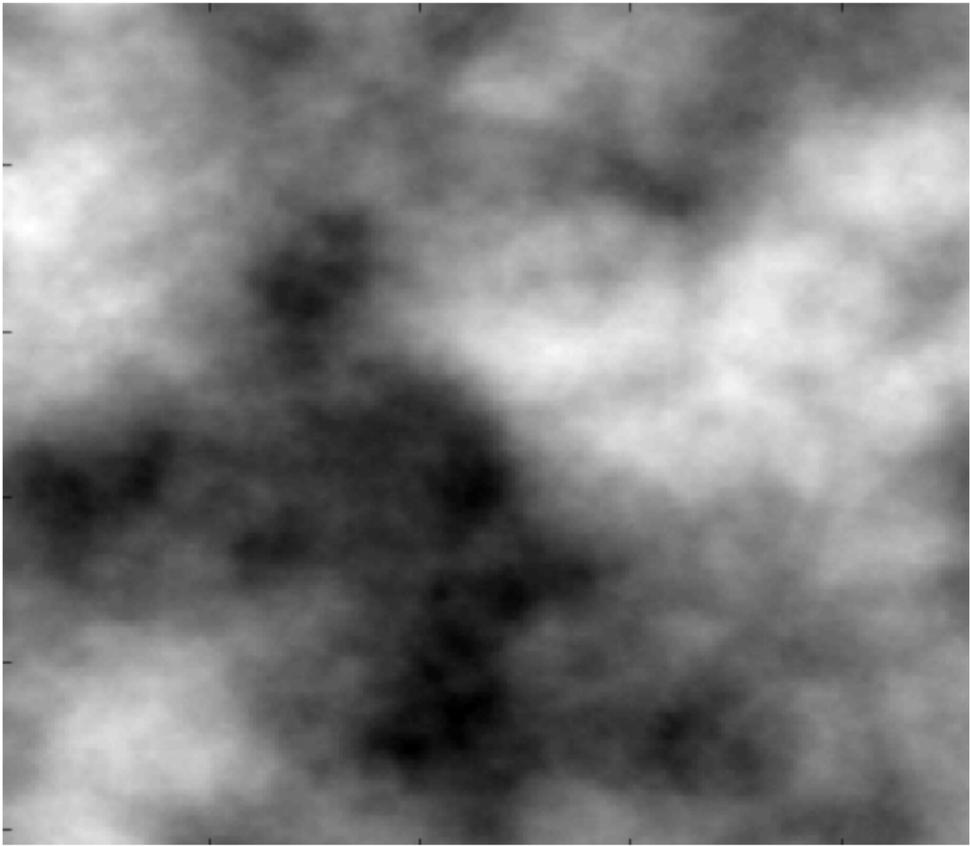
Extrait de [Huang 2000]

- Le modèle gaussien
- Champs markoviens
- Marginales des coefficients d'ondelettes et modèles de dépendance
- Des modèles génératifs, obtenus par combinaisons de formes géométriques (spot noise, booléen, feuilles-mortes, cascades multiplicatives, etc.)
- Modèles génératifs par patchs
- Modèles génératifs neuronaux

- Soit une image I de taille $N \times N$.
- On suppose que $Pr(I) = Z \exp(-\frac{1}{2} I^t \cdot C^{-1} \cdot I)$, avec C matrice de covariance de taille N^4 .
- Si on suppose que I est périodique alors C est diagonalisée par la transformée de Fourier. En notant \hat{I} la transformée de Fourier (discrète) de I on a

$$Pr(I) = Z \exp(-\sum_{\omega} a(\omega) |\hat{I}(\omega)|^2).$$

- La distribution de I ne dépend que de son spectre.



- Le champs I est dit markovien si $I(i, j)$ ne dépend que des valeurs situées dans un voisinage de (i, j) .
- Représentation par champs de Gibbs :

$$Pr(I = x) = \frac{1}{K} \exp \left(- \sum_{c \in C} U_c(x) \right),$$

où C est un système de "cliques", c'est à dire de pixels voisins selon un type de voisinage.

- Exemple : voisinage en 8-connexité et cliques d'ordre 2 ("auto-modèle") :

$$Pr(I = x) = \frac{1}{K} \exp \left(C + a_0 \sum_{(i,j)} x_{(i,j)} + \sum_{k=1}^8 a_k \sum_{i,j} x_{(i,j)} x_{(i,j)+\gamma_k} \right),$$

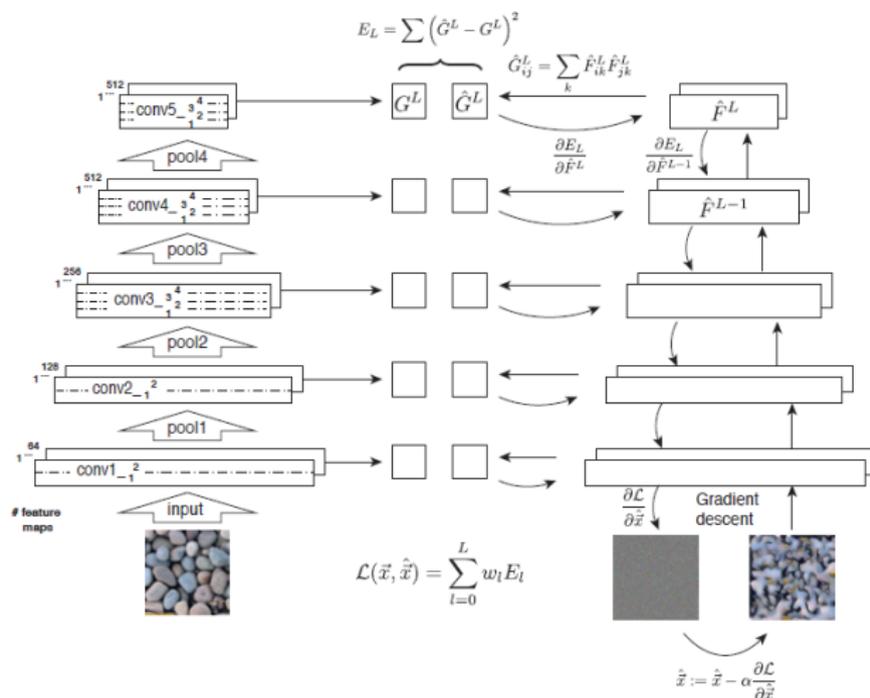
avec $\gamma_k = (0, 1), (1, 0), \dots$

- Problème principal : définition d'énergies multi-échelles (qui respectent un certain nombre d'observations à différentes échelles).

- Marginales relativement simples à modéliser
- Nécessité de prendre en compte les dépendances entre coefficients
- Difficulté à modéliser simplement la présence de bords (représentés par de forts coefficients à toutes les échelles)
- Voir le cours "textures"



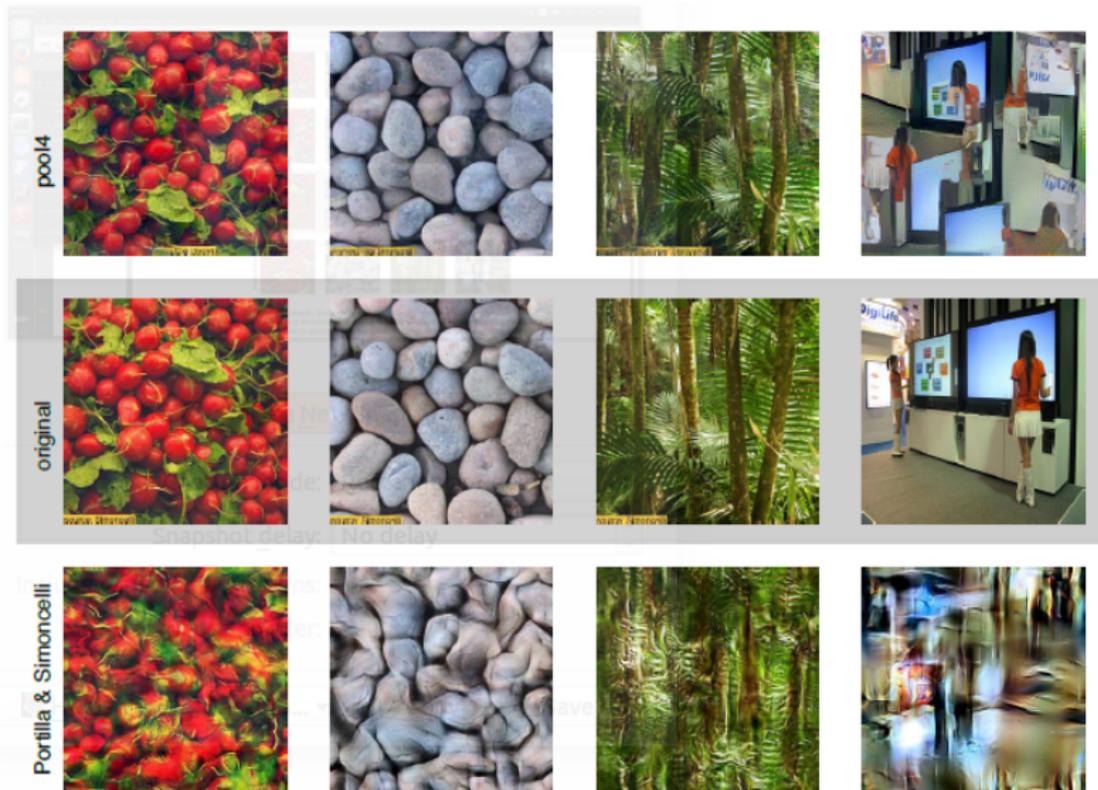
Réseaux profonds convolutionnels ?

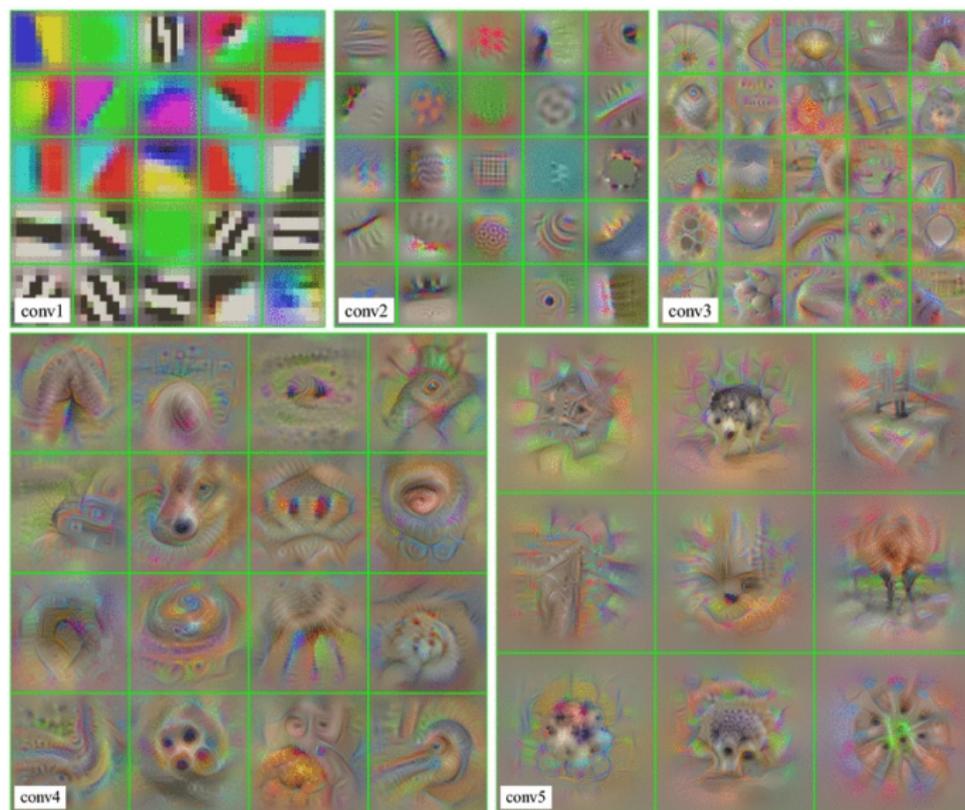


Exemple : Gatys et al. 2015

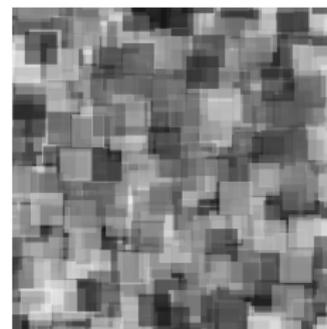
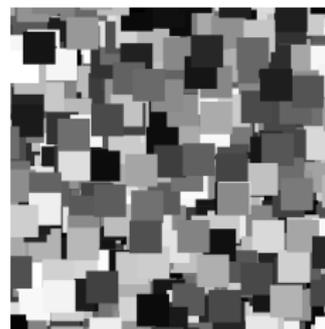
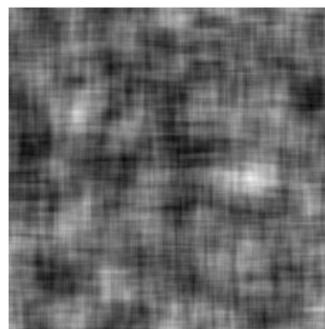
Très nombreuses autres applications des CNN pour la synthèse

Réseaux profonds convolutionnels





Visualisation par maximisation de l'activation (extrait de Mahendran et al. 2016)

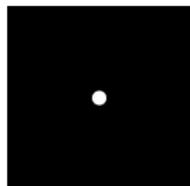


Soit h une fonction de $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ et Φ un processus ponctuel de Poisson stationnaire, d'intensité $\lambda\nu$. Le processus **shot noise** correspondant est défini pour tout $x \in \mathbb{R}^d$ par

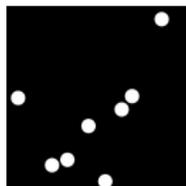
$$f(x) = \sum_i h(x - x_i).$$

(Campbell 1909, Shottky 1918, etc.)

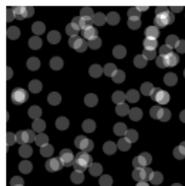
On va s'intéresser au cas $h(x) = \mathbb{1}_X(x)$, où X est un compact de \mathbb{R}^2 .



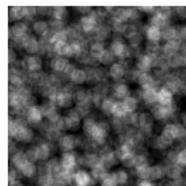
fonction h



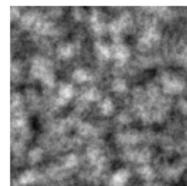
$\lambda = 10$



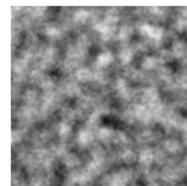
$\lambda = 10^2$



$\lambda = 10^3$



$\lambda = 10^4$



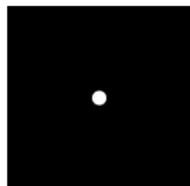
$\lambda = 10^5$

Soit h une fonction de $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ et Φ un processus ponctuel de Poisson stationnaire, d'intensité $\lambda\nu$. Le processus **shot noise** correspondant est défini pour tout $x \in \mathbb{R}^d$ par

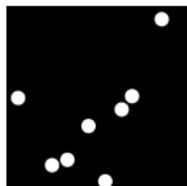
$$f(x) = \sum_i h(x - x_i).$$

(Campbell 1909, Shottky 1918, etc.)

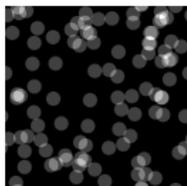
On va s'intéresser au cas $h(x) = \mathbb{1}_X(x)$, où X est un compact de \mathbb{R}^2 .



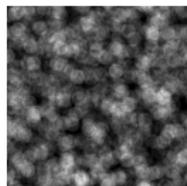
fonction h



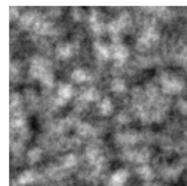
$\lambda = 10$



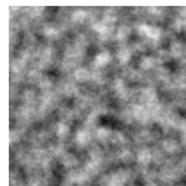
$\lambda = 10^2$



$\lambda = 10^3$



$\lambda = 10^4$



$\lambda = 10^5$

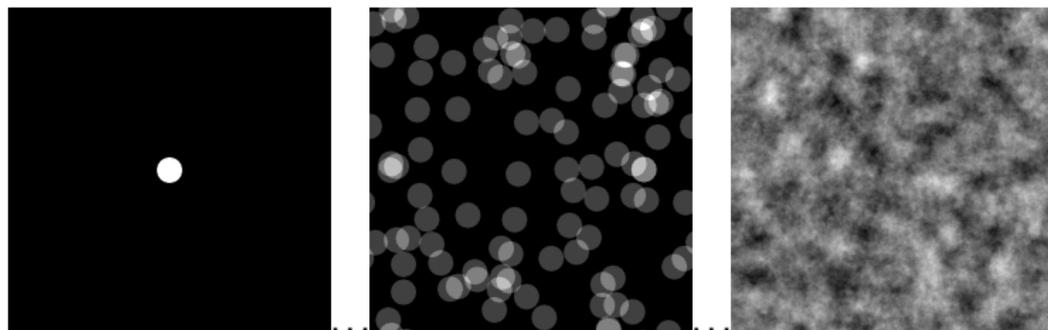
On considère le processus normalisé

$$\tilde{f}(x) = \frac{f(x) - \lambda\nu(X)}{(\lambda\nu(X))^{1/2}}.$$

On montre que si $\lambda \rightarrow +\infty$,

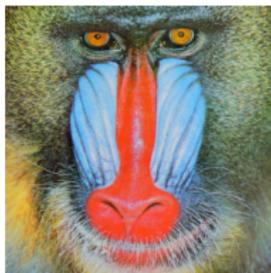
- $\tilde{f}(x) \xrightarrow{\mathcal{L}}$ v.a. gaussienne $N(0, 1)$,
- $\tilde{f} \xrightarrow{\text{fidi}}$ processus gaussien de covariance $C(y) = \nu(X)^{-1}\gamma_X(y)$.

avec ν la mesure de Lebesgue et $\gamma_X(y) = \nu(X \cap (X + y))$.

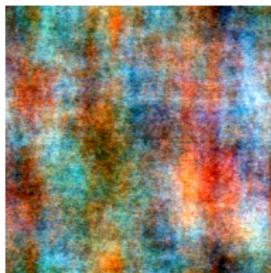


en discret \rightarrow modèles simple de synthèse en couleurs, soit en imposant des phases aléatoires à la TFD, soit en considérant le modèle gaussien limite (Galerne 2011).

Image h



$ADSN$



RPN

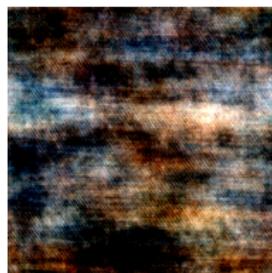
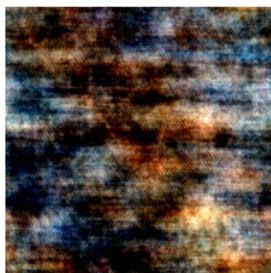
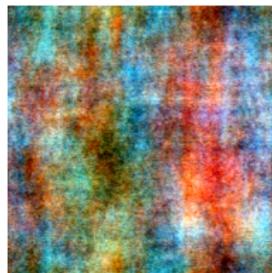
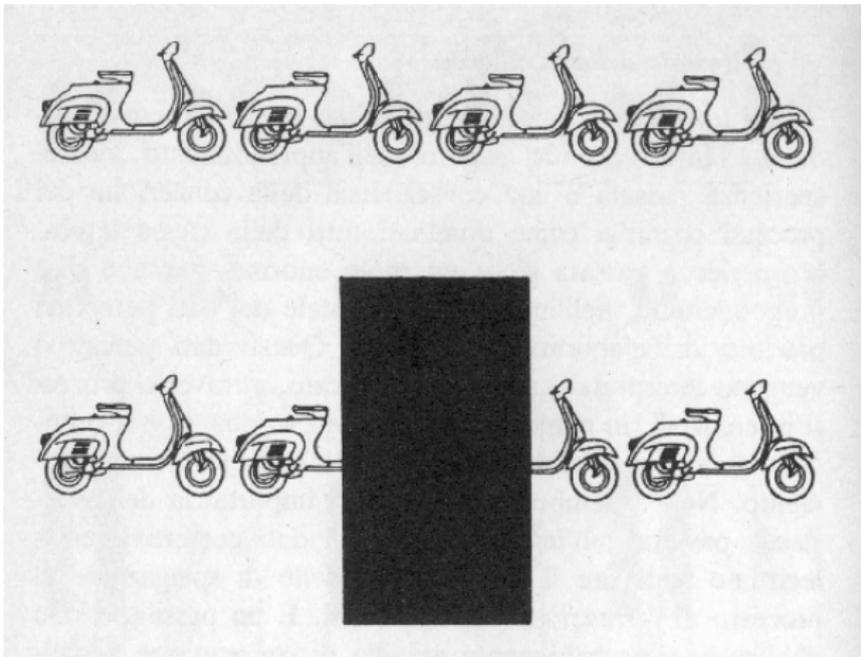


Image h

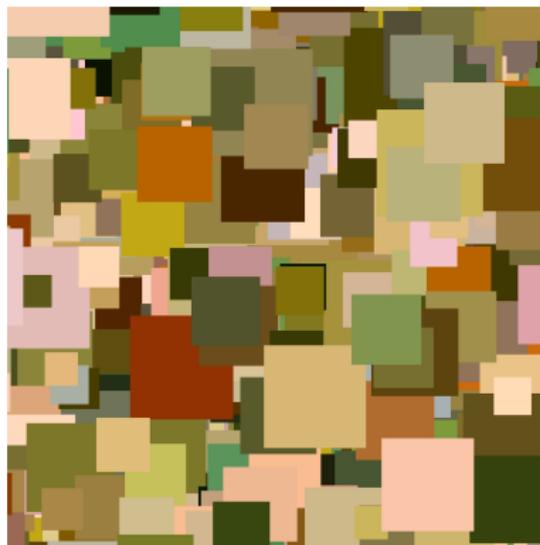
ADSN

RPN





- Modèle aléatoire compatible avec l'occultation
- Obtenu par recouvrements succesifs d'objets indépendants
- Ne dépends que de la distribution d'un ensemble X (aléatoire)



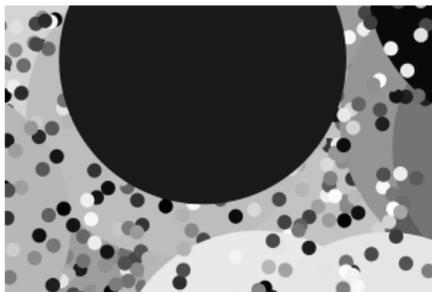
Exemples



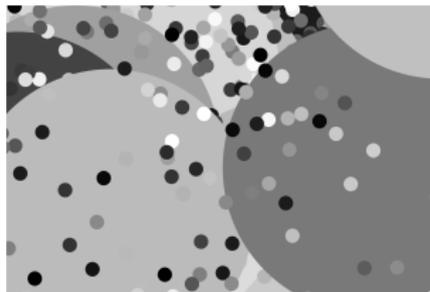
Pour simuler état stationnaire sur un compact Ω de \mathbb{R}^2 : **simulation parfaite** (Propp-Wilson 1996, Kendall-Thönnnes 1999).

Il existe un temps $T < 0$ tel que Ω soit recouvert par les $x_i + X_i$ tels que $t_i \in [T, 0]$.

On simule donc le modèle en superposant des objets jusqu'à avoir recouvert Ω , puis en regardant le recouvrement **par dessous**



Par dessus

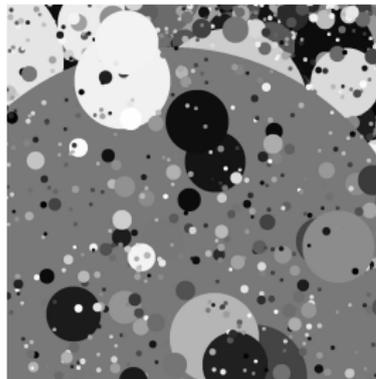


Par dessous

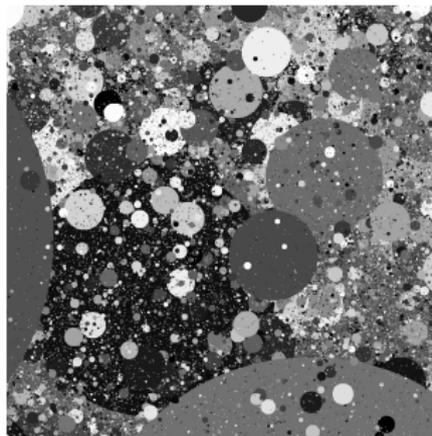
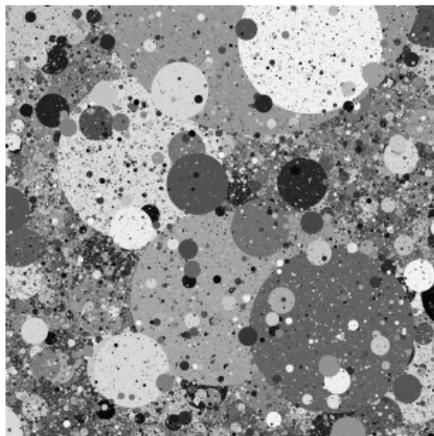
- On choisit des objets X dont la taille suit une loi en puissance : $X = RY$, avec Y un ensemble borné et R une variable aléatoire de densité

$$f(r) = C(r_0, r_1)r^{-\alpha},$$

avec $0 < r_0 \leq r \leq r_1 < +\infty$ et $\alpha > 0$.



- → mêmes observations que sur les images naturelles
- On montre que l'on peut faire tendre r_0 vers 0 et obtenir un modèle qui présente à la fois des bords, des zones homogènes et des lois d'échelles sans limite de détails (Gousseau-Roueff 2007).

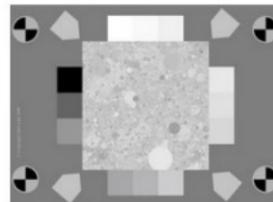


Gauche : $\alpha = 2.9$; Droite : zoom of a factor 10

Compacts et caméraphones appliquent de nombreux traitements (débruitage, réhaussement de contours) → dégradation des textures

Mire Feuilles Mortes

Mire compatible IEEE-CPIQ et ISO pour l'évaluation du bruit et de la préservation des textures. (Également incluse dans les modules Autofocus & Timing, et Photo.)

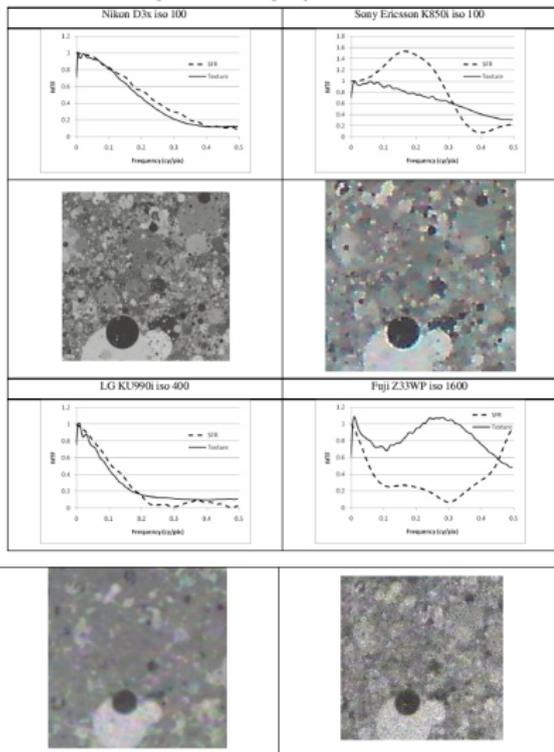


Protocole d'évaluation DxoMark

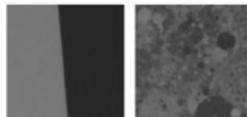
On mesure $\frac{|I|^2}{|I_0|^2}$, où I est l'image capturée et I_0 l'image feuilles-mortes idéale.

Différents comportements

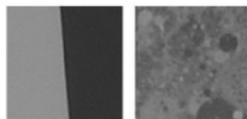
Table 1. The four different configurations of texture vs. edge sharpness.



A detail of a target made of a dead leaves pattern, designed to measure Texture Acutance. It is obtained by drawing random shapes that occlude each other in the plane, like dead leaves falling from a tree. The statistics of this model follow the distribution of the same statistics in natural images.



In this example from a DSLR with edge enhancement, sharpness seems equal on edge and on texture. Many details are visible in the texture.



In this second example edges have been digitally enhanced, and the edge looks over-sharp, with visible processing halos ("ringing"). On the texture part, many details have disappeared.

Le même modèle (avec distribution de couleurs adéquate) peut être utilisé pour entraîner un réseau de neurones profonds :

- de restauration (débruitage, super-résolution) (Achddou et al. 2022)
- de reconnaissance (Baradad et al. 2021)

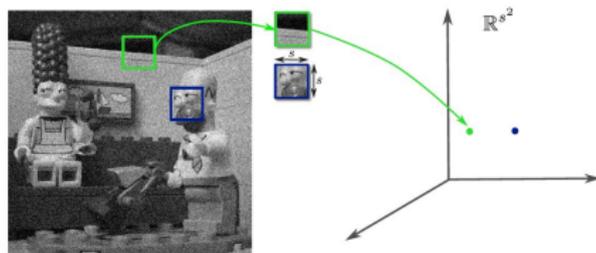
→ Qq paramètres versus bases volumineuses

→ Permet de s'affranchir de bases spécifiques et de s'adapter à des modalités particulière (e.g. acquisition d'images)



Fig. 3: Denoising comparison with different FFDNet trainings. Top row from left to right : clean image, noisy image with $\sigma = 25$, model trained on white noise, model trained on Gaussian fields. Bottom row from left to right : model trained on dead leaves images with fixed radius $r = 100$, model trained on the dead leaves dataset, model trained on the mixed dataset, model trained on natural images.

Une modélisation globale de la distribution des pixels : hors de portée
→ modèles sur des imagerie, les *patches* (cf cours de restauration)



Illus. Julie Delon

- Sparsité (DCT, Yu-Sapiro 2011, dictionnaires adaptatifs, Elad-Aharon 2006)
- Mélange de gaussiennes : Yu et al. 2010 (PLE), Zoran-Weiss 2011 (EPLL), Houdard et al. 2018, etc.
- Modèles locaux gaussien (Lebrun et al. 2013, NL-Bayes)
- Mélange de gaussiennes généralisées (Deledalle et al. 2018)

Trois grandes familles d'architectures profondes pour capturer la distribution des images naturelles :

- Réseaux antagonistes génératifs (GAN, Goodfellow et al. 2014)
- Auto-encodeurs (e.g. ae variationnels (VAE, Kingma-Welling 2014)
- Modèles de diffusion (Sohl-Dickstein et al. 2015, Ho et al. 2020)

→ Exemples vus dans les cours précédents de restauration, synthèse de texture, inpainting.

→ Suite dans les cours du S2 :

- Modèles génératifs pour l'image (Galerie, Leclaire)
- Apprentissage profond pour la synthèse et la restauration d'images (Almansa, Ladjal, Newson)