

Analyse et synthèse de textures

Yann GOUSSEAU

Télécom ParisTech - LTCI CNRS

Master IMA - Cours TDI
2013-2014

1 Généralités

2 Analyse

- Statistiques d'ordre deux et cooccurrences
- Champs de Gibbs
- Filtres de Gabor, ondelettes et variantes

3 Synthèse

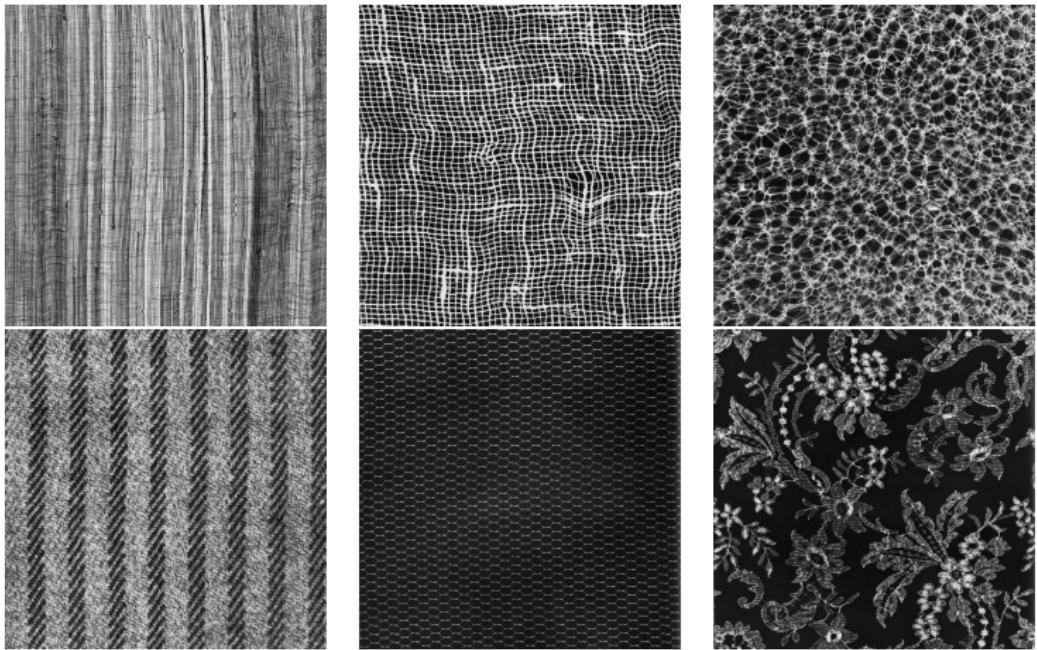
- Quelques modèles ad'hoc
- Méthodes par rééchantillonnage

Définition d'une texture ?

- **Textile** : disposition et mode d'entrecroisement des fils dans un tissage.
- **Culinaire** : Qualité physique des aliments liée à leur densité, leur viscosité, leur caractère homogène, leur dureté.
- **Matériaux** : Caractéristique relative à la dimension, la forme, la disposition des grains dans le métal.
- **Arts** : texture musicale, texture d'un drame.

Même si l'on se restreint à la vision : pas de définition complètement satisfaisante.

- Zone homogène en un certain sens, assemblage d'entités élémentaires formant un tout.
- Description des variations d'intensité lumineuse sur une surface, rendant compte de propriétés telles que la rugosité, la douceur, la granularité, etc.



Textures de l'album de Brodatz

Analyse de texture

- Difficultés

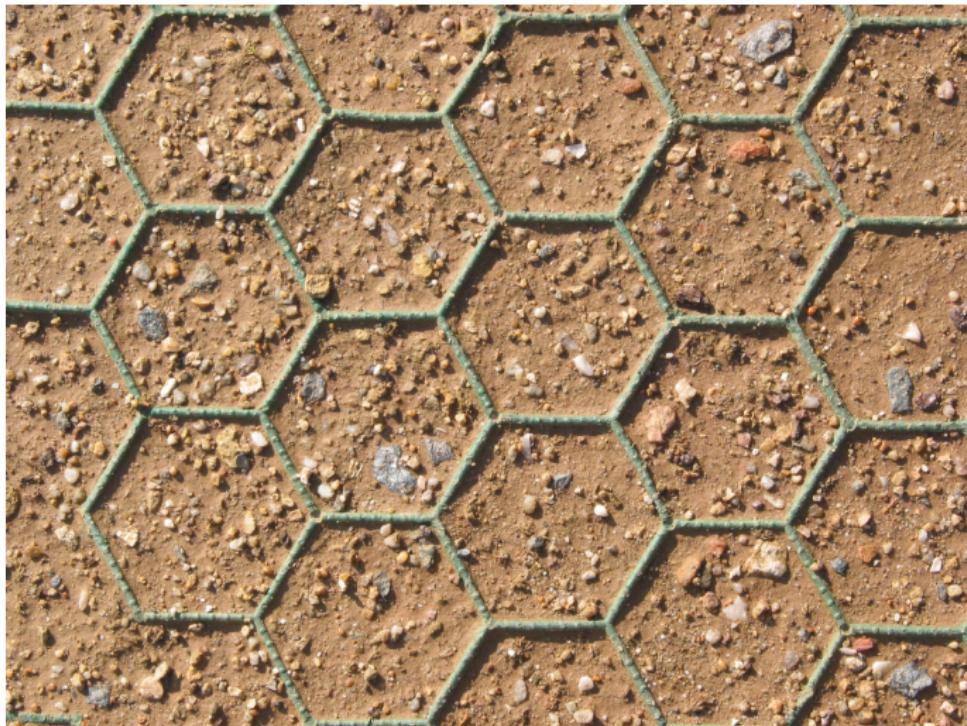
- Pas de modèle mathématique suffisement général et manipulable
- Pas de notion satisfaisante de distance entre textures
- Notion fortement dépendante de l'échelle

- Applications

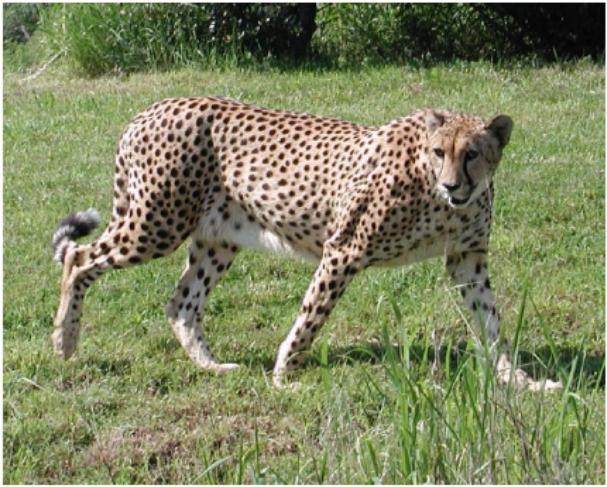
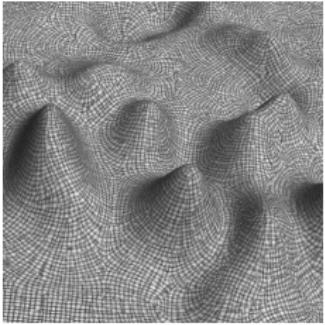
- Reconnaissance et classification de matériaux, contrôle de qualité.
- Reconnaissance d'objets, segmentation.
- Reconstruction de volume (“shape from texture”).
- Synthèse et restauration d'images.











Texture et forme

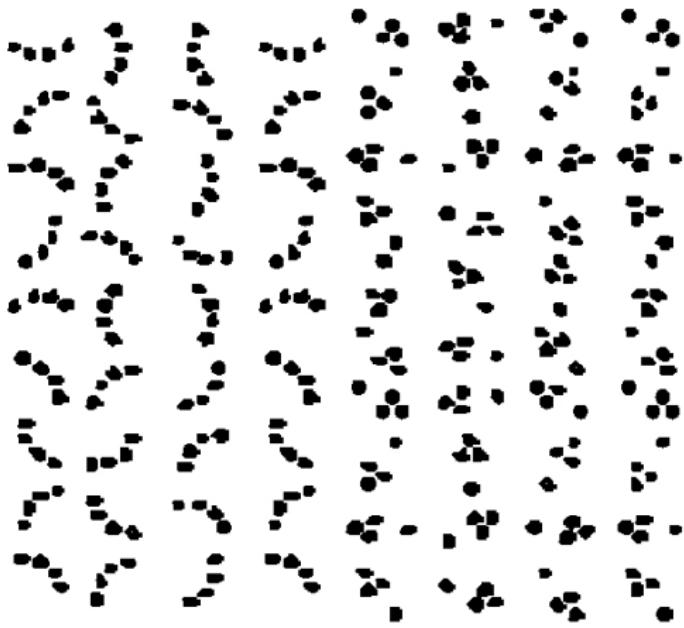


La plupart des modélisations considèrent une texture comme la réalisation d'un champ aléatoire.

Problème : quelles sont les statistiques pertinentes pour la perception d'une texture ?

- Beck 73 : statistiques d'ordre 1 sur des "propriétés" (contraste, orientation, couleurs)
- Julesz 73-75 : statistiques d'ordre 1 et 2 entre valeurs ponctuelles de niveau de gris
- Julesz 81 : statistiques d'ordre 1 et 2 entre "textons" (éléments structurés, terminaisons, jonctions en "T",...)

Problème : à quels objets mathématiques correspondent les textons ?



Stationnarité

- Un champ aléatoire $Z : \mathbb{Z}^2 \rightarrow \mathbb{R}$ est dit *stationnaire* si pour tout $n \in \mathbb{N}$, $z \in \mathbb{Z}^2$ et $z_1, \dots, z_n \in \mathbb{Z}^2$
 $(Z(z_1), \dots, Z(z_n))$ et $(Z(z_1 + z), \dots, Z(z_n + z))$
ont même distribution jointe.
- Les textures sont généralement supposées stationnaires.
- En pratique, on dispose d'une image I définie sur un domaine limité Ω .
 - on peut considérer un champ périodique associé à l'image
 - ou une notion de stationnarité réduite aux déplacements inclus dans Ω

Les statistiques du premier ordre

On considère une image I définie sur $\Omega = [1, \dots, N]^2$ à valeurs dans $[1, \dots, G]$. Pour $g \in [1, \dots, G]$, on note $f(g)$ la fréquence empirique du niveau g

$$f(g) = \frac{1}{N^2} \sum_{z \in \Omega} 1_{I(z)=g}.$$

Parmi les statistiques fréquemment utilisées pour décrire une texture :

- Moyenne $\mu(I) = \frac{1}{N^2} \sum_z I(z)$,
- Variance $\sigma^2(I) = \frac{1}{N^2} \sum_z (I(z) - \mu)^2$
- Energie $E(I) = \frac{1}{N^2} \sum_z I(z)^2$,
- Entropy $Ent(I) = - \sum_{g=1}^G f(g) \log(f(g))$, etc.

Statistiques du premier ordre

Limitation : toutes ces quantités ne dépendent que du seul histogramme de l'image.

Elles prennent des valeurs arbitraires sous l'effet d'un *changement de contraste*, c'est à dire une transformation

$$I \mapsto h \circ I,$$

avec h une fonction réelle croissante.

Remarques sur l'estimation

- En général : on ne dispose que d'une seule réalisation du champ
→ l'estimation se fait par moyennage sur des zones suffisement grandes
- Dans le cas de scènes complexes : plusieurs zones ayant des textures différentes : l'hypothèse de stationnarité est fausse.
On estime les paramètres sur des voisinages bornés ("fenêtres glissantes").
Compromis précision de l'estimation - localisation



Statistiques du second ordre : la covariance

- La covariance d'un champ stationnaire Z (de moyenne nulle) est définie par $C(z) = E(Z(0)Z(z))$.
- La covariance de l'image I est estimée, pour tout $z_1 \in \Omega$, par

$$C(z_1) = \frac{1}{N^2} \sum_{z \in \Omega} I(z)I(z + z_1).$$

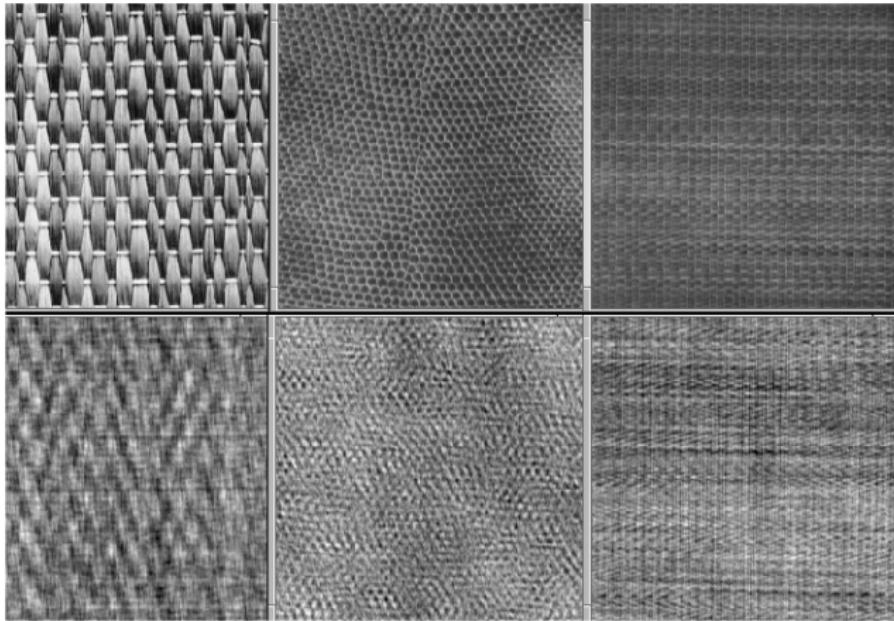
formule où l'on suppose I périodique.

- Théorème de Wiener-Khintchin : si \hat{Z} désigne la transformée de Fourier de Z , alors

$$|\hat{Z}(\omega)|^2 = \hat{C}(\omega).$$

La covariance

- Une large classe de textures ("micro-textures") sont bien représentées par leur covariance (donc par le module de la transformée de Fourier, le *spectre* de l'image).
- Plusieurs méthodes ont été proposées pour extraire l'information pertinente du spectre (e.g. Tang-Stewart 2000, Wang-Knoesen 2007).
- Pour tester cette assertion : méthode de synthèse.
 - On calcule la TFD de l'image I , $\{\hat{I}_{i,j}\}$,
 - Pour chaque i,j soit $\hat{J}_{i,j} = |\hat{I}_{i,j}| \exp(i\phi_{i,j})$, où $\phi_{i,j}$ est uniformément distribué entre 0 et 2π ,
 - J est l'image dont la TFD est $\{\hat{J}_{i,j}\}$.



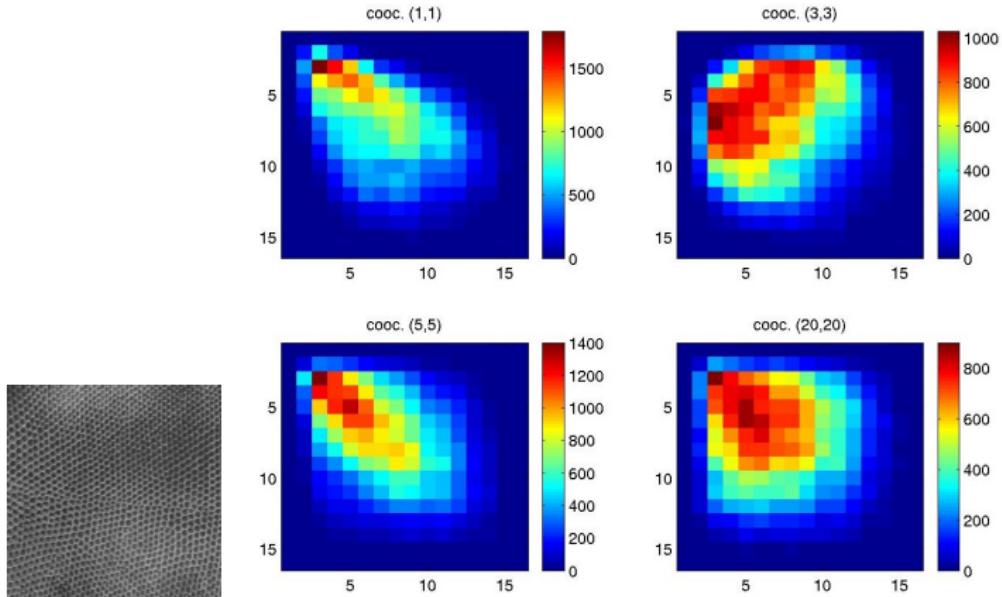
Matrices de cooccurrences (Haralick et al, 1973)

- Pour $(k, l) \in [1, \dots, N]^2$ et $(a, b) \in [1, \dots, G]^2$, la matrice de cooccurrence $M_{k,l}$ de I est définie par

$$M_{k,l}(a, b) = \frac{1}{(N-k)(N-l)} \sum_{i=1}^{N-k} \sum_{j=1}^{N-l} \mathbf{1}(I(i,j) = a \text{ et } I(i+k, j+l) = b).$$

- Matrices de taille $G \times G$
- Contiennent toutes les statistiques du second ordre de l'image
- On considère généralement (k, l) petits devant la taille de l'image, et on quantifie l'image sur un nombre de niveaux $G' < G$ (typiquement 16).
- Pour → information locale, les matrices sont calculées sur des fenêtres.

Exemples de matrices de cooccurrences



Matrices de cooccurrences

Pour réduire la dimension des attributs retenus, on considère des fonctionnelles sur les matrices de cooccurrence.

- Energie $E(k, l) = \sum_{a,b} M_{k,l}(a, b)^2$ (homogénéité),
- Entropie $Ent(k, l) = -\sum_{a,b} M_{k,l}(a, b) \log(M_{k,l}(a, b))$,
- Inertie $I(k, l) = \sum_{a,b} (a - b)^2 M_{k,l}(a, b)$ (dispersion), etc.

Eventuellement → invariance par rotation en moyennant sur plusieurs orientations.

Champs de Markov

- Le champ Z est dit markovien si $Z(i,j)$ ne dépend que des valeurs situées dans un voisinage de (i,j) (cf cours de F. Tupin).
- Les textures sont supposées markoviennes et peuvent être représentées par des champs de Gibbs, c'est à dire que

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

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

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

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

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

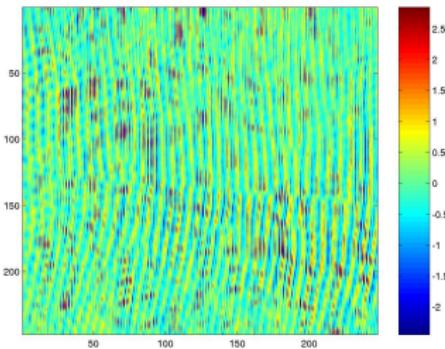
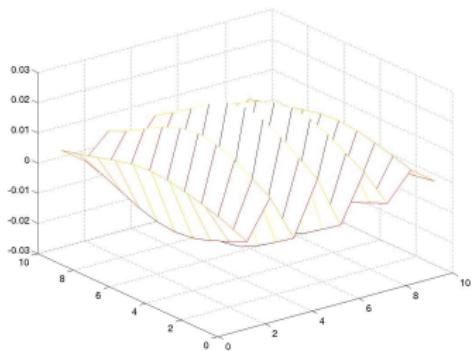
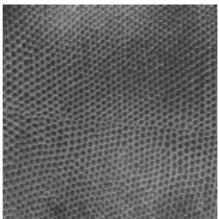
- Problèmes principaux : trouver l'ordre du modèle (des cliques) et les potentiels U_c .

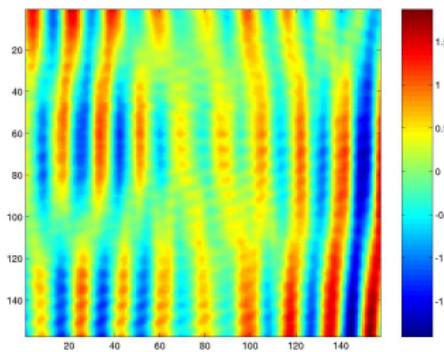
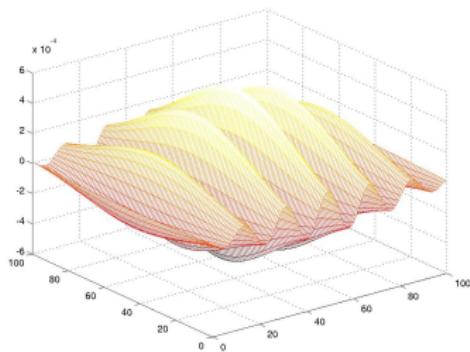
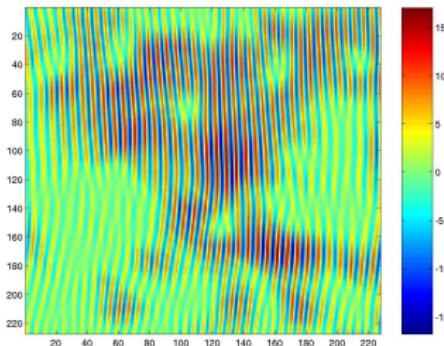
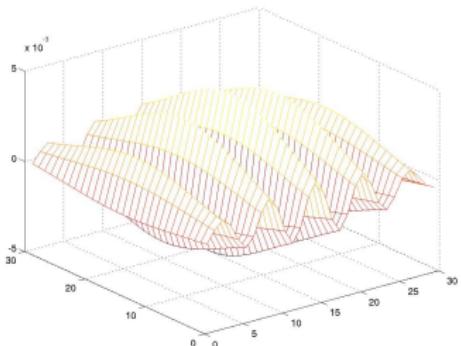
Filtres de Gabor

- Les textures sont représentées par leurs réponses à des filtres linéaires extrayant une information fréquentielle localisée et orientée.
- Filtres de Gabor horizontal en $(0,0)$ de paramètres $(\sigma_1, \sigma_2, f_0)$:

$$\phi(x, y) = e^{-\frac{1}{2} \left(\frac{x^2}{\sigma_1^2} + \frac{y^2}{\sigma_2^2} \right)} \cos(2\pi f_0 x).$$

- Les autres filtres s'obtiennent par rotations de θ et translations de (x_0, y_0) .
- La texture est caractérisée par différentes statistiques sur les réponses aux filtres (énergie).
- Ordres de grandeur : 5-10 orientations et 5-10 fréquences.





Filtres de Gabor et réponses

Décomposition sur des bases d'ondelettes

- On utilise en général des trames d'ondelettes redondantes, comportant plus de 3 orientations et invariantes par translation.
- Comme précédemment on peut retenir l'énergie comprise dans chaque bande fréquentielle.
- Une solution alternative : modélisation paramétrique des coefficients. Par exemple, si $c_{k,l,\eta}$ désigne le coefficient à la position k , échelle l et orientation η , on peut supposer qu'à l et η fixés, les $\{c_{k,l,\eta}\}_k$ sont i.i.d. selon

$$c_{k,l,\eta} \sim Ce^{-\left(\frac{|x|}{\beta}\right)^\alpha}.$$

- Les textures sont alors représentées par les coefficients α et β (pour chaque l et η).

Synthèse de textures

- **But** : partant d'une image de texture A , synthétiser des images ponctuellement différentes de A , mais qui produisent la même impression visuelle.
- **Applications** :
 - Graphisme
 - Amélioration, modification d'images
 - Psychophysique
 - Test de la pertinence des modèles
- **Difficultés** :
 - Pas de modèle mathématique complètement satisfaisant des textures
 - Pas de critère quantitatif d'évaluation des résultats



Disparition du commissaire Yezhov (image www.newseum.org)

Quelques méthodes ad'hoc

- Stockage d'une image de texture, puis plaquage de la texture avec des conditions aux bords → problèmes topologiques.
- Perturbation des normales (Blinn 78) :

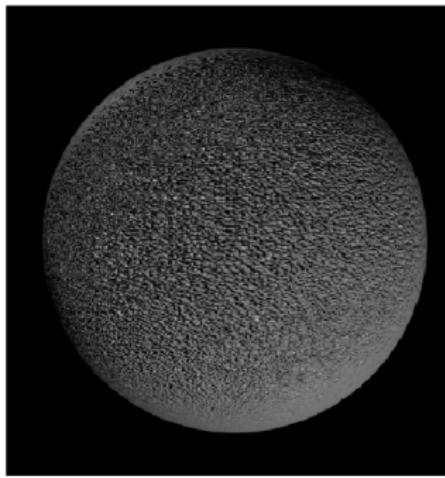
à partir d'une surface, on calcule en chaque point une direction normale fictive, après perturbation de la surface, puis on utilise cette nouvelle normale dans un modèle de reflection :

$$\mathbf{M}(x, y) \rightarrow \mathbf{M}(x, y) + F(x, y)\mathbf{n}(x, y),$$

ce qui implique

$$\mathbf{n} \rightarrow \mathbf{n} + \frac{\partial \mathbf{M}}{\partial x} \wedge \mathbf{n} \frac{\partial F}{\partial y} - \frac{\partial \mathbf{M}}{\partial y} \wedge \mathbf{n} \frac{\partial F}{\partial x}$$

c'est donc la fonction F qui modélise physiquement la texture : table de valeurs, valeurs aléatoires filtrées ...



Perturbation des normales d'une sphère

- Texture solide

chaque point de \mathbb{R}^3 reçoit une valeur (couleur), puis la texture est obtenue par une coupe dans cet espace : textures de type marbre, bois, etc...

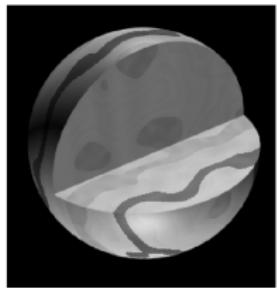
- Reaction-diffusion :

modèle initialement proposé par Turing (1952) pour les motifs apparaissant sur la peau de certains animaux. Un motif est le résultat de la concentration en morphogènes, qui apparaissent, diffusent, et disparaissent au cours du temps.

$$\frac{\partial C}{\partial t} = a^2 \Delta C - bC + R.$$

Pour modéliser l'anisotropie, ΔC devient

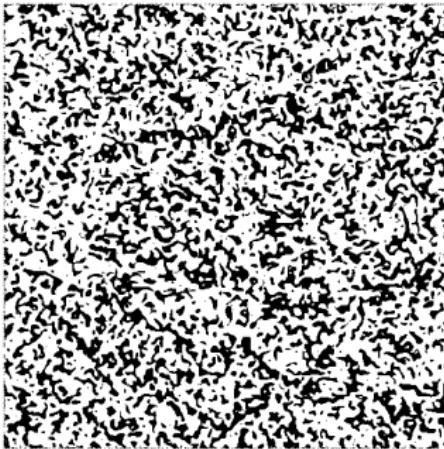
$$a_1 \frac{\partial^2 C}{\partial x^2} + a_2 \frac{\partial^2 C}{\partial x \partial y} + a_3 \frac{\partial^2 C}{\partial y^2}$$



Texture solide



All the same, but different



Reaction-diffusion

Modèles “fractals”

Essentiellement des simulations de généralisations du mouvement brownien : les mouvements browniens fractionnaires (Mandelbrott-Van Ness 68).

- dimension 1 : V_H processus tel que

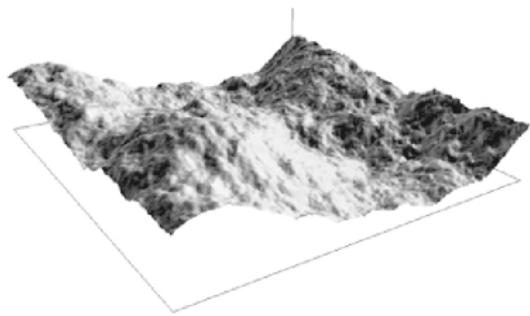
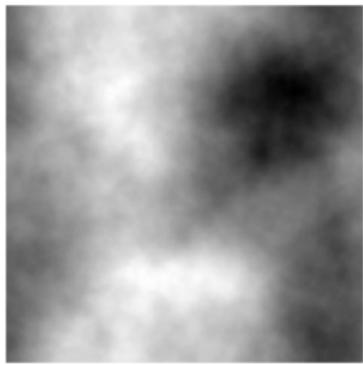
$$V_H(t_1) - V_H(t_2) \sim N(0, |t_2 - t_1|^{2H}).$$

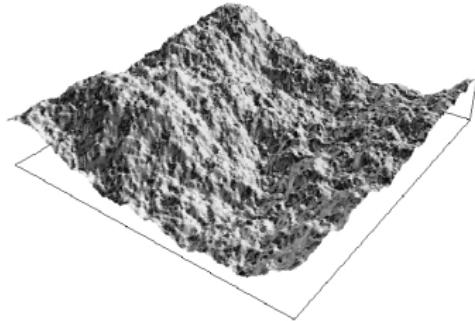
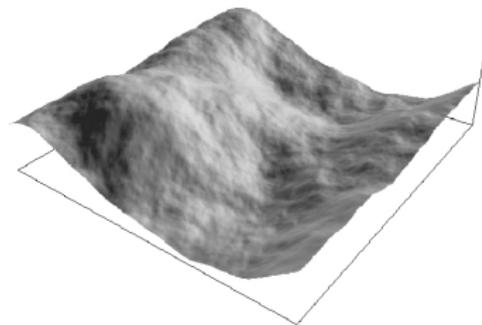
Remarque :

$V_H(t_0) - V_H(t_0 + t)$ et $r^{2-H} (V_H(t_0) - V_H(t_0 + rt))$ ont alors même distribution.

- dimension 2 : $V_H : \mathbb{R}^2 \rightarrow \mathbb{R}$ tel que

$$V_H(x_1) - V_H(x_2) \sim N(0, ||x_1 - x_2||^{2H}).$$



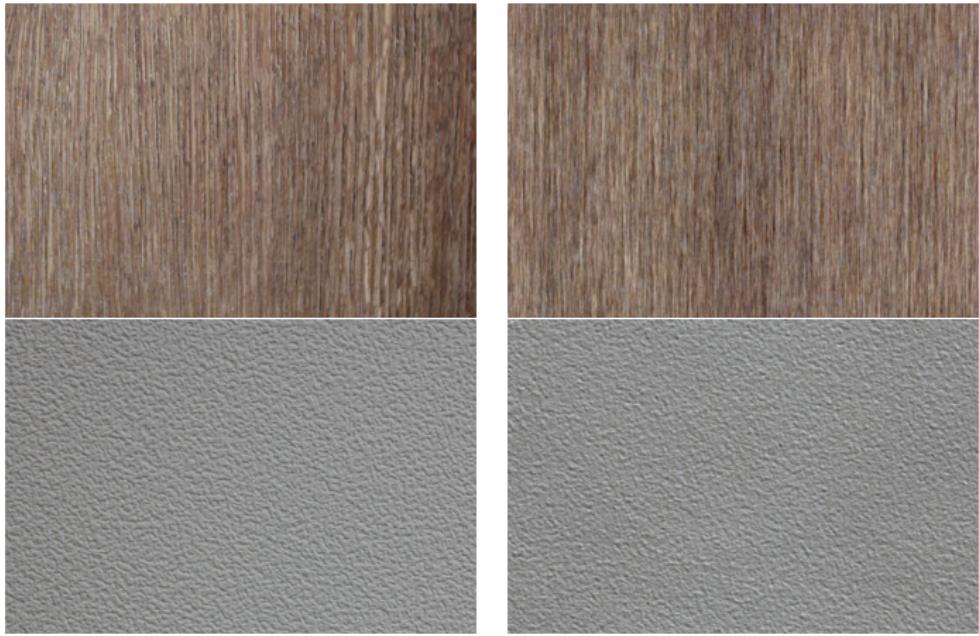


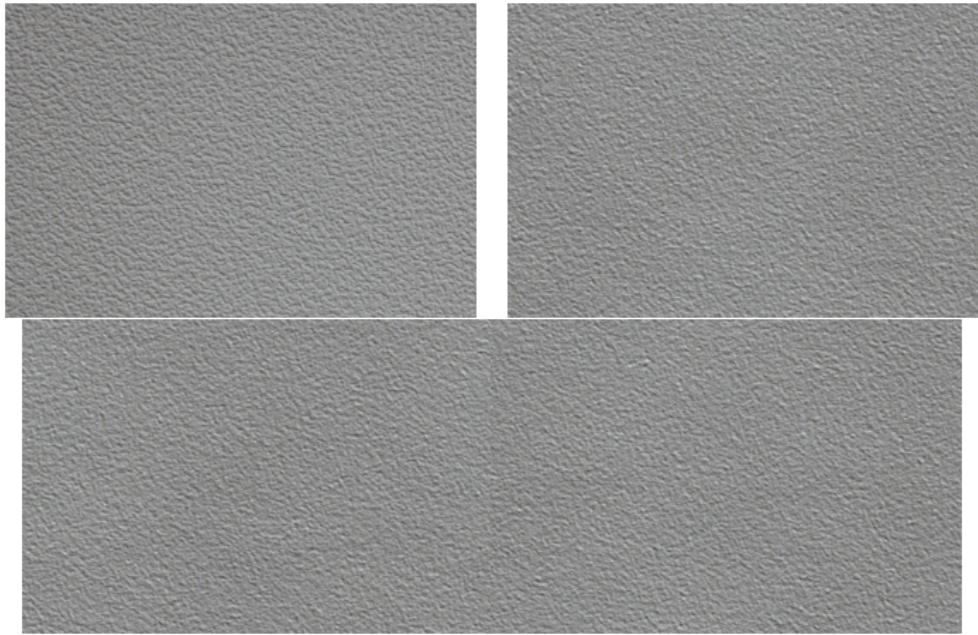
Deux modèles génériques

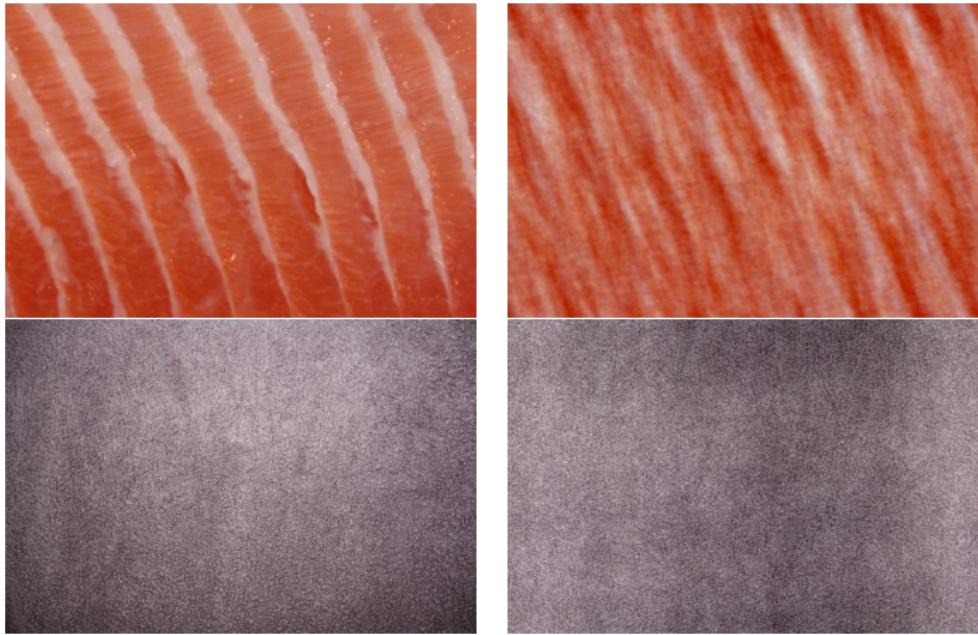
Synthèse par phases aléatoires (Galerne-Gousseau-Morel 2011)

- La forme de base est remplacée par une image couleur
- Deux options pour la synthèse :
 - On impose des phases aléatoires
 - On synthétise le modèle limite gaussien dont on connaît la covariance

→ résultats visuellement similaires
- On impose une cohérence de phase entre canaux couleurs
- Suppression des artefacts de non-périodicité
- → Démo en ligne ! (*random phase demo* sous google)









Plusieurs régimes pour une même texture physique

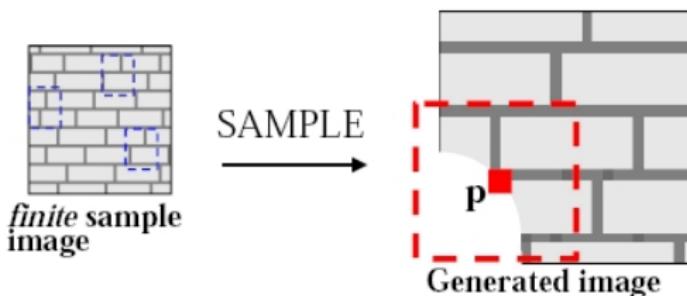


Les méthodes de rééchantillonnage

(Efros-Leung '99)

Idée :

- considérer la texture comme la réalisation d'un **champs de Markov**
- au lieu d'échantillonner à partir d'un modèle du type Gibbs (dont les paramètres sont appris sur la texture), on **échantillonne directement sur une image exemple A**
- la synthèse s'effectue pixel par pixel
- pour chaque nouveau pixel de la zone à synthétiser, on suppose qu'une partie de son voisinage est déjà synthétisé, et **on cherche dans A le pixel dont le voisinage est le plus proche.**



(fig : Efros-Leung 99)

- Soient Ω_1 et Ω_2 les domaines de A et B respectivement (sous-ensembles de \mathbf{Z}^2)
- Soit $F = \{a_1, \dots, a_K\}$ un sous-ensemble de \mathbf{Z}^2 .
- Pour $p \in \Omega_1$, soit $N_A(p) = \{A(p + a_1), \dots, A(p + a_K)\}$, le voisinage de p correspondant à F .
- On définit une distance entre voisinages, e.g. :

$$d_v(N_A(p), N_B(p')) = \frac{1}{K} \sum (A(p + a_i) - B(p' + a_i))^2$$

- pour un pixel $p' \in \Omega_2$, et $\epsilon > 0$ soit

$$O_\epsilon(p') = \{p \in \Omega_1 | d_v(N_A(p), N_B(p')) < \epsilon\}$$

Pour synthétiser une nouvelle texture B : pour chaque $p' \in \Omega_2$

- On calcule $O_\epsilon(p')$,
- on tire un pixel p de $O_\epsilon(p')$ uniformément,
- on affecte la valeur $B(p') = A(p)$

... mais bien sûr, on ne sait pas calculer $O_\epsilon(p')$, puisque les valeurs prises par B dans le voisinage de p' ne sont pas connues.

La distance d est alors modifiée pour ne prendre en compte que les pixels de B où la synthèse a déjà été effectuée.

Puis la synthèse s'effectue séquentiellement, en partant d'un petit morceau de l'image originale A (quelques pixels).

Avantages : Excellent résultats, même avec des textures très structurées.

Inconvénients : Très lent, très sensible au choix de B , instable.

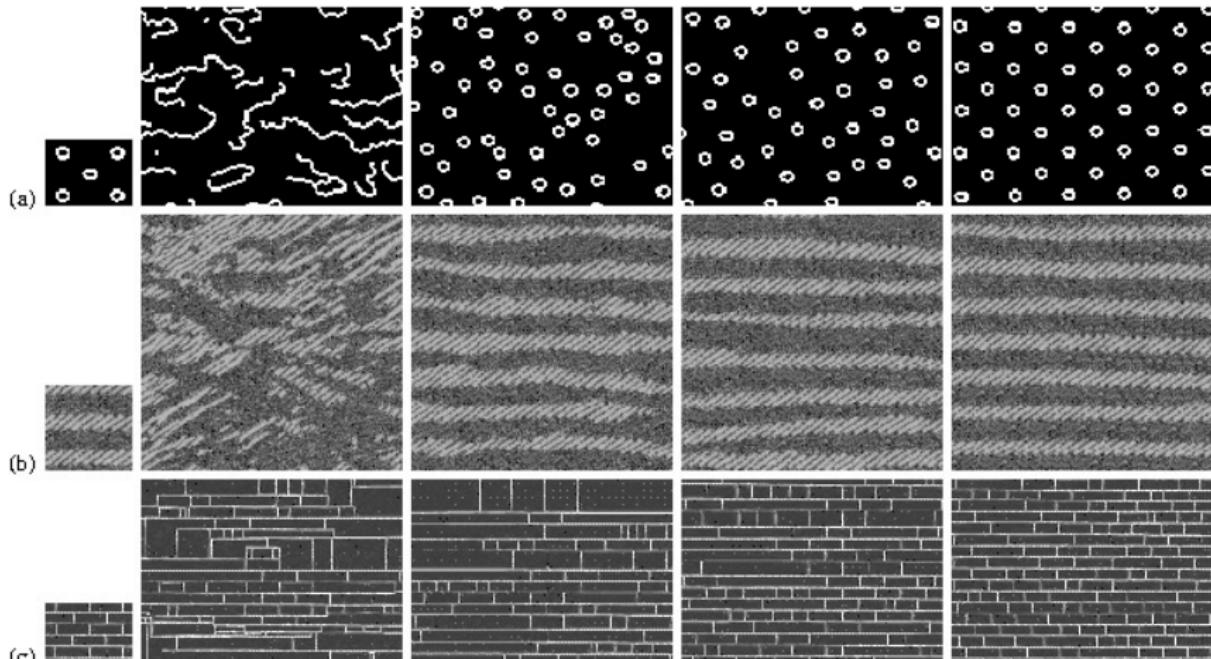
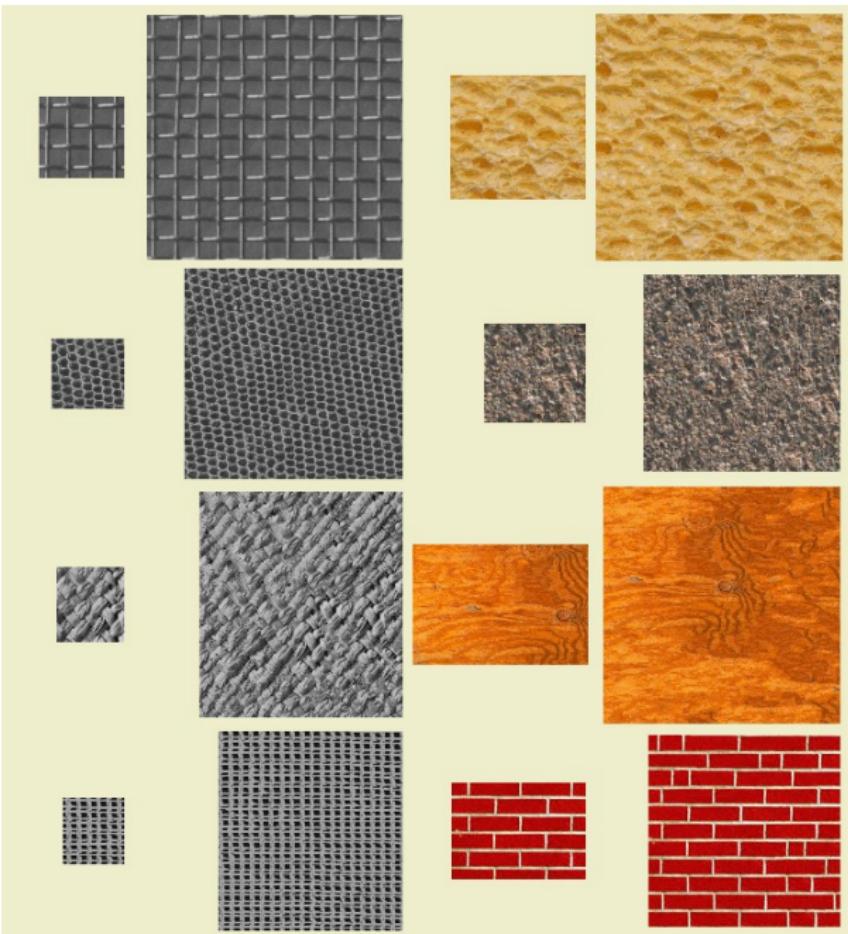


Figure 2. Results: given a sample image (left), the algorithm synthesized four new images with neighborhood windows of width 5, 11, 15, and 23 pixels respectively. Notice how perceptually intuitively the window size corresponds to the degree of randomness in the resulting textures. Input images are: (a) synthetic rings, (b) Brodatz texture D11, (c) brick wall.

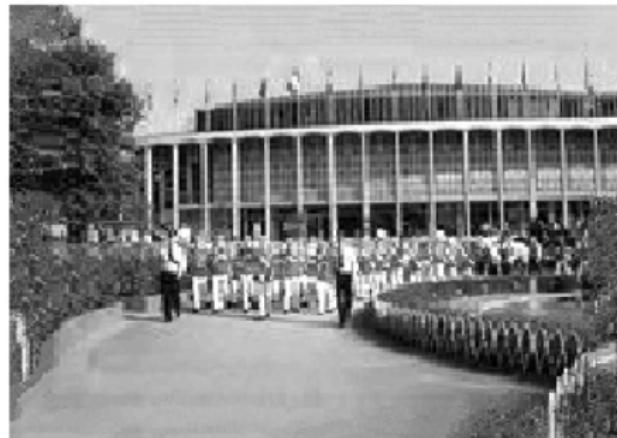
Extrait de Efros-Leung 1999



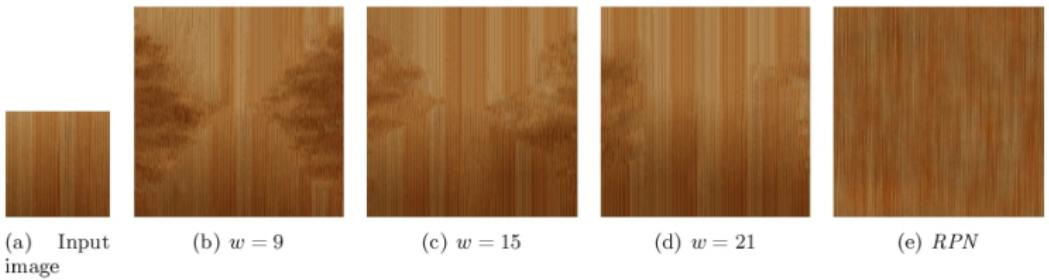
ut it becomes harder to lau-
ound itself, at "this daily-
wing zooms," as House Der-
escribed it last fall. He fail-
ut he left a ringing question
more years of Monica Lewin-
inda Tripp?" That now seem
Political comedian Al Frat-
extphase of the story will

de reme a the ral counta rself, at "this da Lew acrue y-
it ndatyears counve Tring rooms," as Heft he fast nd it l-
ars dat noears ortseas ribed it last nt hest bedian Al. I-
economical Homd it h Al. Heft ars Gf, as da Lewindailf J-
lian Al Ths," as Lewing questies last aticarsticall. He-
is dian Al last fal counda Lew, at "this dailyears d ily-
edianicall. Hoozewing rooms," as House De fale f De-
und itical counoestscribed it last fall. He fall. Hefft
rs oroheneon it nd it he left a ringing questica Lewin-
icars coecoms," astore years of Monica Lewinow seee
a Thas Fring zoome stooniscat nowea re left a rouuse
bouestof Mhe lelfta Lest fast nging läumesticars Hef-
nd it rip?" Trhousef, a ringind it'sonestid it a ring que-
astical cois ore years of Moung fall. He ribof Mouse
ore years ofanda Tripp?" That hedian Al Lest fassee yea-
nda Tripp? Tpolitical comedian Aléthe fw se ring que-
olitical cona re years of the storears oafas l Frat nica L-
ros Lew se lesta rime l He fas quest nging of, at beou

Extrait de Efros-Leung 1999



Extrait de Efros-Leung 1999



(a) Input
image

(b) $w = 9$

(c) $w = 15$

(d) $w = 21$

(e) RPN

Input Image



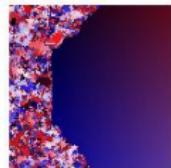
Output Image



Position Map



Copy Map



Autres méthodes de rééchantillonnage

- Même principe, mais on échantillonne sur la décomposition en ondelettes de la texture de départ. Le voisinage pris en compte pour la synthèse est à la fois spatial et entre échelles.
- Les pixels sont remplacés par des groupes structurés de pixels (ensembles de niveau, fenêtre avec des conditions de bord,...). La synthèse s'effectue par agencement des formes (e.g. Pérez-Gangnet-Blake 2004)
- Respect simultané de la géométrie et de la texture
- Nombreuses variantes
(Etat de l'art : Wei-Lefebvre-Kwatra-Turk 2009)