

Analyse et synthèse de textures

Yann Gousseau

Télécom Paris - Institut Polytechnique de Paris

MVA
2024-2025

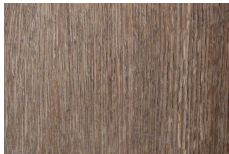
1 Généralités

2 Analyse

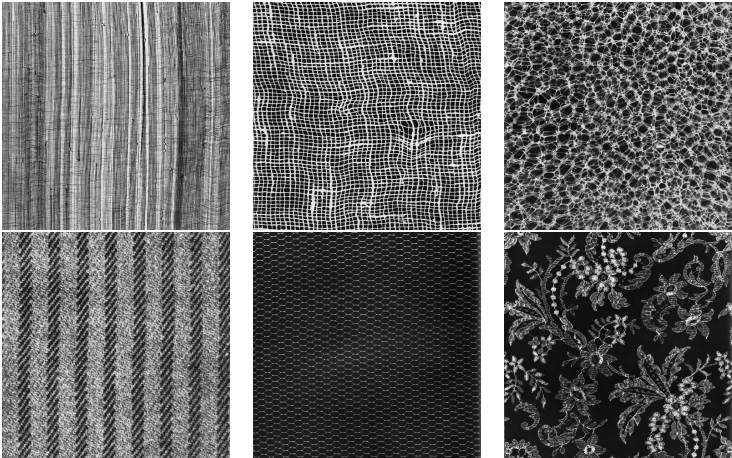
- Statistiques pixéliques
ordre deux et cooccurrences ; champs de Gibbs
- Approche structurelle et morphologie mathématique
- Filtres de Gabor, ondelettes et variantes
- Descriptions invariantes :
descripteurs locaux versus apprentissage

3 Synthèse

- Quelques modèles ad'hoc
- Méthodes par contraintes statistiques :
valeurs ponctuelles, ondelettes, dictionnaires parcimonieux, CNN
- Méthodes par copier-coller de patches



Quelques textures



Textures de l'album de Brodatz

De nombreuses significations

- **Matériaux** : Caractéristique relative à la dimension, la forme, la disposition des grains dans un métal,
- **Culinaire** : Qualité physique des aliments liée à leur densité, leur viscosité, leur caractère homogène, leur dureté,
- **Textile** : disposition et mode d'entrecroisement des fils dans un tissage,
- **Infographie** : apparence donnée à une surface (généralement par une image),
- Autres définitions en cristallographie, cosmologie, musique, peinture, etc.

Définitions possibles dans le cadre de la vision

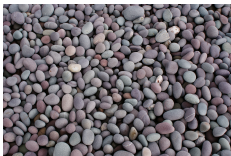
- Zone homogène d'une image, constituée de la répétition d'entités élémentaires formant un tout.
« Un subtile équilibre entre répétition et innovation »
Yves Meyer
- Description des variations d'intensité lumineuse sur une surface, rendant compte de propriétés telles que la rugosité, la douceur, la granularité, etc.

De nombreuses significations

- **Matériaux** : Caractéristique relative à la dimension, la forme, la disposition des grains dans un métal,
- **Culinaire** : Qualité physique des aliments liée à leur densité, leur viscosité, leur caractère homogène, leur dureté,
- **Textile** : disposition et mode d'entrecroisement des fils dans un tissage,
- **Infographie** : apparence donnée à une surface (généralement par une image),
- Autres définitions en cristallographie, cosmologie, musique, peinture, etc.

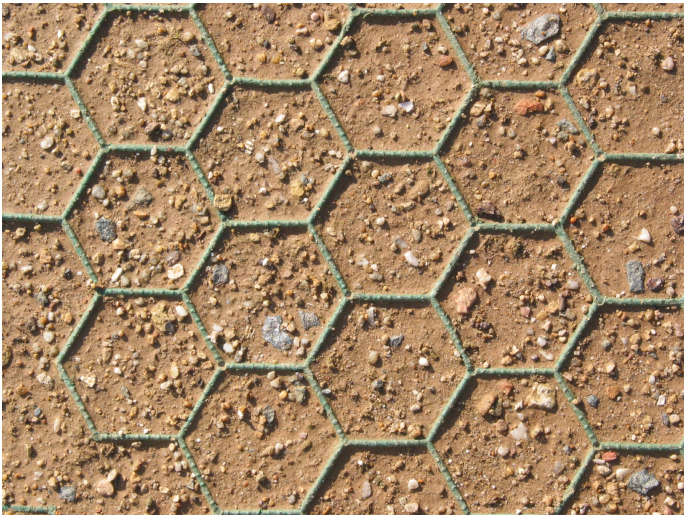
Définitions possibles dans le cadre de la vision

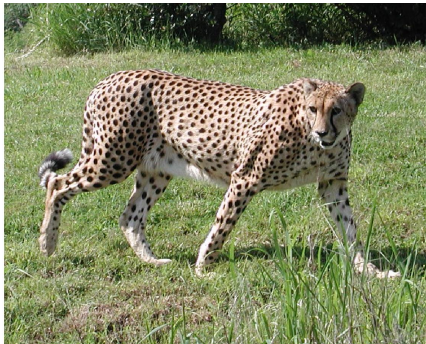
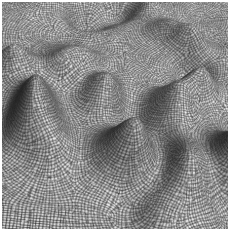
- Zone homogène d'une image, constituée de la répétition d'entités élémentaires formant un tout.
« Un subtile équilibre entre répétition et innovation »
Yves Meyer
- Description des variations d'intensité lumineuse sur une surface, rendant compte de propriétés telles que la rugosité, la douceur, la granularité, etc.









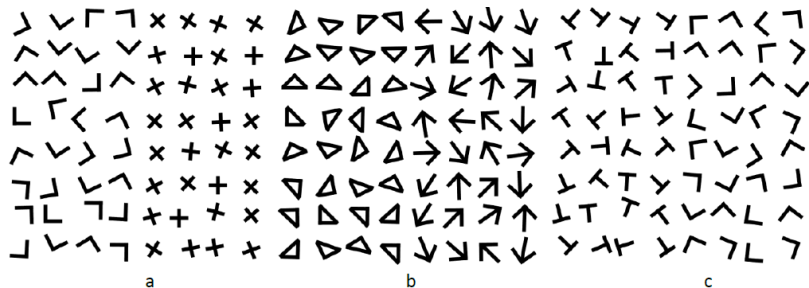


Texture et forme



- Modèle simple de texture : répétition de textons identiques (et souvent circulaires) ;
- L'apparence des textons dépend de la géométrie de la surface ;

Notre système visuel a la capacité de grouper des éléments en régions homogènes en quelques dixièmes de secondes (vision pré-attentive)



From Rosenholtz 2014

Problème : quelles sont les statistiques pertinentes pour expliquer la discrimination ?¹

- [Beck 1973] : statistiques d'ordre 1 sur des "propriétés" (contraste, orientation, couleurs)
- [Julesz 1973-75] : statistiques d'ordre 1 et 2 entre valeurs ponctuelles de niveau de gris

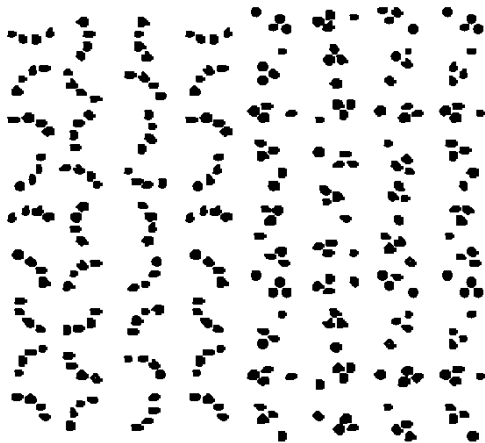
Exemples de paires indiscernables ayant même stats d'ordre 2 et différentes stats d'ordre 3 ou 4

- [Julesz 1981] : statistiques d'ordre 1 et 2 entre "textons" (éléments structurés, terminaisons, jonctions en "T",...)

Contre-exemple de textures discriminable ayant même stats d'ordre 2 (e.g. slide suivant)

Problème : à quels objets mathématiques correspondent les textons ?

1. Remarque : c'est l'utilisation des ordinateurs qui a permis de vraiment attaquer ce problème. En 1950, F. Attneave étudie la perception d'un bruit blanc ... dessiné à la main avec un générateur de nombres aléatoires !



Contre-exemple à la première conjecture de Julesz

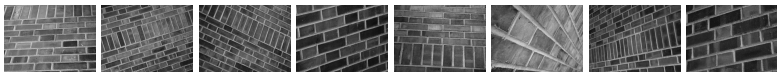
Première partie : analyse de texture

But :

$$I \rightarrow \{a_1(I), \dots, a_n(I)\}$$

où les a_i sont

- aussi compatibles que possible avec notre perception,
- éventuellement invariants à certaines transformations (géométriques et radiométriques)



Echantillons de textures de la base UIUCTex (Lazebnik *et al.* 2005)

Première partie : analyse de texture

● Difficultés

- La pertinence perceptuelle d'un modèle est une notion subjective et difficile à quantifier
Pas de définition objective d'une distance entre textures
- Grande variété d'images de textures / difficulté à définir un modèle générique
- Notion fortement dépendante de l'échelle

● Applications

- Reconnaissance et classification de matériaux, contrôle de qualité
- Reconnaissance d'objets, segmentation
- Synthèse d'images
- Edition d'images (inpainting)
- Restauration d'images
- Reconstruction de volume ("shape from texture")

Première partie : analyse de texture

- **Difficultés**

- La pertinence perceptuelle d'un modèle est une notion subjective et difficile à quantifier
Pas de définition objective d'une distance entre textures
- Grande variété d'images de textures / difficulté à définir un modèle générique
- Notion fortement dépendante de l'échelle

- **Applications**

- Reconnaissance et classification de matériaux, contrôle de qualité
- Reconnaissance d'objets, segmentation
- Synthèse d'images
- Edition d'images (inpainting)
- Restauration d'images
- Reconstruction de volume ("shape from texture")

Stationnarité

- On représente une texture par un champs aléatoire *stationnaire*
- 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.
- 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 $[1, \dots, G]$.

Un estimateur de la loi marginale du champs est son **histogramme**

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

pour $g \in [1, \dots, G]$.

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 (voir le cours sur la radiométrie).

Néanmoins : la distribution des couleurs joue un rôle important dans notre perception des textures

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 suffisamment grandes (hypothèse d'ergodicité)
- Dans le cas de scènes complexes : plusieurs zones ayant des textures différentes : l'hypothèse de stationnarité est fautive.
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 $z \in \Omega$, par

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

formule où l'on suppose I périodique.

- Théorème de Wiener-Khintchin (version discrète) : si \hat{I} désigne la transformée de Fourier discrète de I , alors

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

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. Wang 1990, Tang-Stewart 2000, Wang-Knoesen 2007, Florindo-Bruno 2012, etc.).

- Pour tester cette assertion : **méthode de synthèse**.

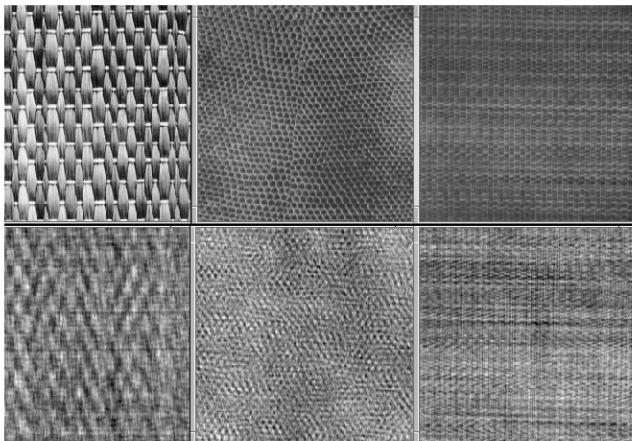
Partant d'une image exemple I :

- Soit $\hat{I} = TFD(I)$,
- 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 ,

- On synthétise $J = TFD^{-1}(\hat{J})$.



Haut : images originales
Bas : synthèses

La suite dans la partie du cours sur la synthèse

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

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

- Estimateur de

$$\mathbb{P}(I(z) = a, I(z + (k, l)) = b)$$

(caractérise complètement le second ordre)

- Matrices de taille $G \times G$
- 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 \rightarrow information locale, les matrices sont calculées sur des fenêtres.

Méthode ancienne et élémentaire

Bon compromis complexité / caractère discriminant.

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

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

- Estimateur de

$$\mathbb{P}(I(z) = a, I(z + (k, l)) = b)$$

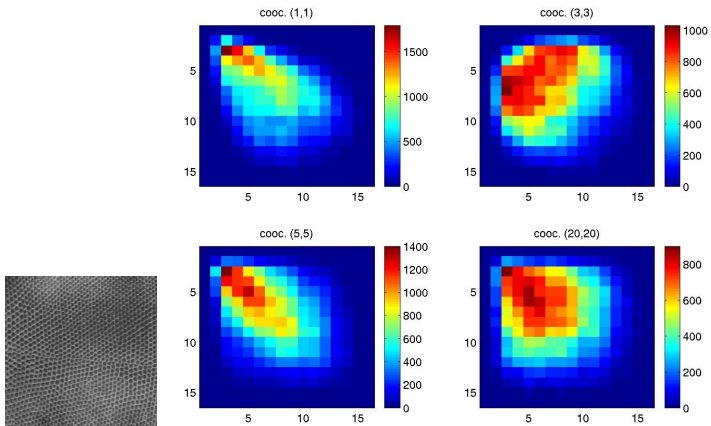
(caractérise complètement le second ordre)

- Matrices de taille $G \times G$
- 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 \rightarrow information locale, les matrices sont calculées sur des fenêtres.

Méthode ancienne et élémentaire

Bon compromis complexité / caractère discriminant.

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 \rightarrow invariance par rotation en moyennant sur plusieurs orientations.

Champs de Markov

- Le champ Z est dit markovien si la distribution de $Z(z)$ est complètement caractérisée par ses voisins (sur un voisinage prescrit).
- 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-connexité 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 la forme des potentiels U_c .
 - développer des modèles multi-échelles

Morphologie mathématique

- Outils d'analyse non-linéaires introduits dans les années 1960 à l'Ecole des Mines (G. Matheron, J. Serra), pour l'analyse des milieux poreux
- Idée de départ : sonder des structures à l'aide d'ensembles géométriques simples (éléments structurants)
- Outils géométriques efficaces pour l'analyse d'images et en particulier de textures

Un exemple d'opération : l'ouverture.

Soit deux ensembles X et B (élément structurant) . L'ouverture de X par B est

$$O_B(X) = \bigcup_{x \in \mathbb{R}^2 : x+B \subset X} (x+B)$$

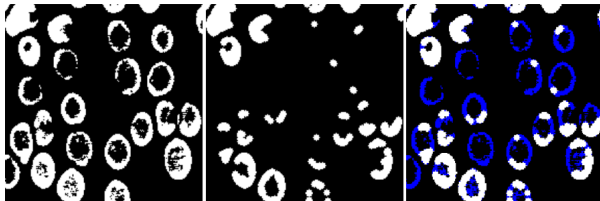


Image Isabelle Bloch

Outil d'analyse de textures : les granulométries

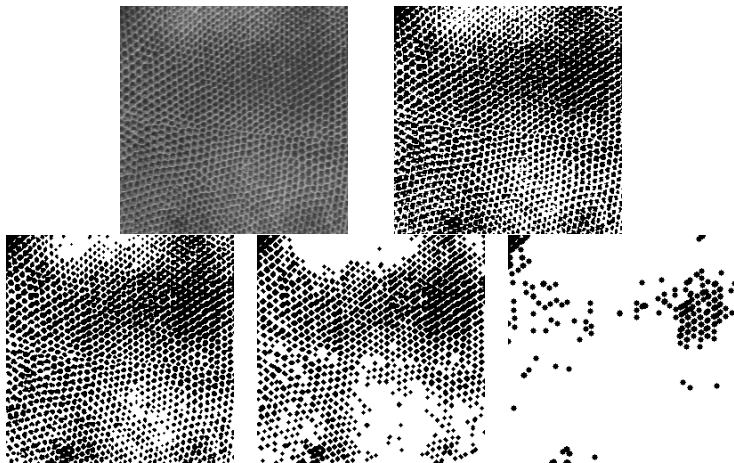
- Pour une texture binaire I on considère la suite des ouvertures $O_{rA}(I)$, où les rA sont des éléments structurants dilatés du facteur r .
- La courbe de granulométrie de I est le graphe de la fonction

$$r \rightarrow \nu(\mathbb{1}(O_{rA}(I) = 1))$$

où $\nu(A)$ est la surface de A .

- On caractérise généralement la texture par la dérivée des courbes de granulométrie pour différents éléments structurants.
- Pour une texture à niveau de gris, on considère différents ensembles de niveau

$$\chi_\lambda(I) = \{(i,j) : I(i,j) \geq \lambda\}.$$



Filtres de Gabor

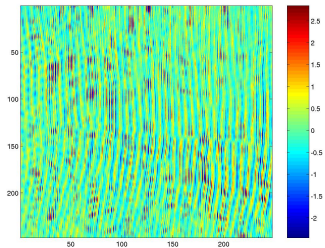
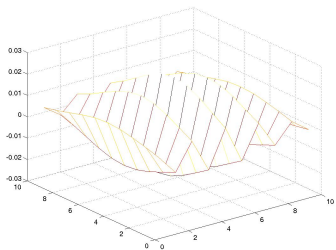
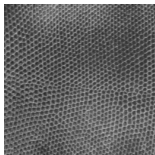
- Les textures sont représentées par leurs réponses à des filtres linéaires extrayant une information fréquentielle localisée et orientée.
- Filtres de Gabor horizontaux :

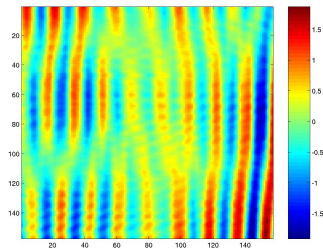
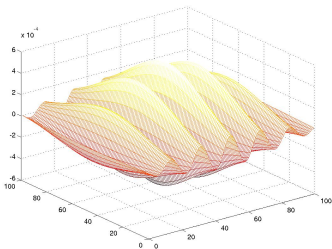
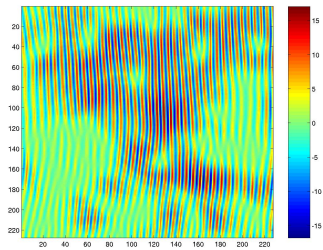
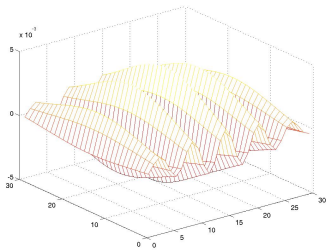
$$\phi(x, y) = e^{-\frac{1}{2}\left(\frac{x^2+y^2}{\sigma}\right)} \cos(2\pi fx),$$

$$\phi(x, y) = e^{-\frac{1}{2}\left(\frac{x^2+y^2}{\sigma}\right)} \sin(2\pi fx),$$

- Les autres filtres s'obtiennent par rotations de θ .
- La texture est caractérisée par différentes statistiques sur les réponses aux filtres (classiquement : l'énergie).
- Ordres de grandeur : 5-10 orientations et 5-10 fréquences.
- Variante : gaussienne non-isotrope

Très nombreuses variantes avec d'autres représentations espace-fréquence





Factres de Gabor et r ponses

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 distribués selon

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

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

Analyse de textures et invariances

- Représentation des textures

$$I \rightarrow \{a_1(I), \dots, a_n(I)\}$$

- En fonction de l'application, on peut imposer des invariances
 - géométriques

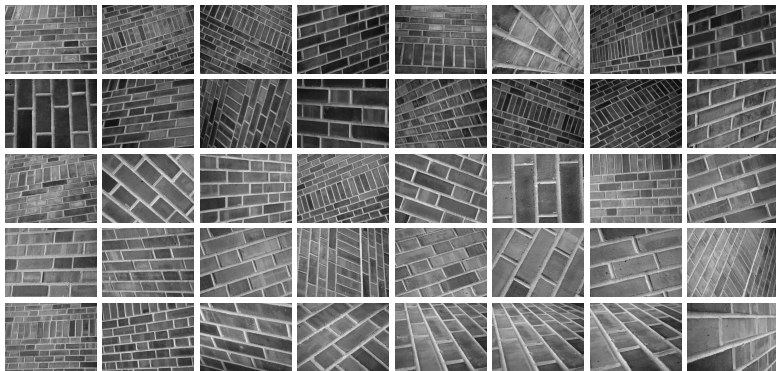
$$a_i(I \circ T) = a_i(I),$$

où T est une transformation géométrique (rotation, changement d'échelle, transformation affine, transformation projective, etc.)

- radiométriques

$$a_i(f \circ I) = a_i(I),$$

où f est une fonction croissante (changement de contraste).

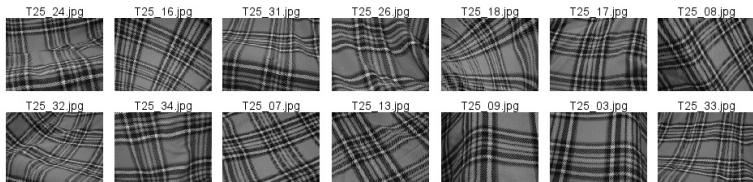
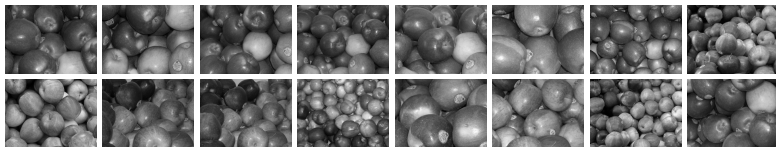


Echantillons de textures de la base UIUCTex (Lazebnik *et al.* 2005)

Analyse de texture et invariances

- Statistiques ponctuelles (covariances, co-occurrences, champs de Markov)
 - Rotations : moyennes des réponses à différentes orientations
 - Autres transformations (en particulier changements d'échelle) : difficile
- Filtrage (Gabor, ondelettes)
 - Transformations géométriques (rotations, échelle) : utilisation de familles redondantes
 - Changements de contraste linéaires : normalisation des réponses

Textures non-planes, non-rigides



- Invariance affine ou projective insuffisante
- Approches possibles :
 - Apprentissage (Varma 2002, Cimpoi 2015, etc.)
 - Invariance aux déformations géométriques (similitudes ou affinités) **locales** (Lazebnik et al. 05, Mellor et al. 08, Xia et al. 10, Nguyen et al. 2011)

Utilisation des descripteurs SIFTs (Lazebnik et al. 2005)

- Extraction de points clés et de descripteurs SIFTs
- Construction d'un dictionnaire par classification non-supervisée (k-means)
- Calcul des fréquences d'utilisation des mots (sacs de mots)

L'invariance à des transformation complexe est obtenue par invariances à des transformations plus simples (similitudes, affines) **localement**.

Utilisation de la carte topographique pour l'analyse de texture (Xia et al. 2010)

- Extraction de formes S (composantes connexes d'ensembles de niveau)
- Représentation : distribution d'attributs de formes invariants
- Attributs : obtenus à partir des moments

$$\mu_{pq}(s) = \int \int_s (x - \bar{x})^p (y - \bar{y})^q dx dy,$$

Puis :

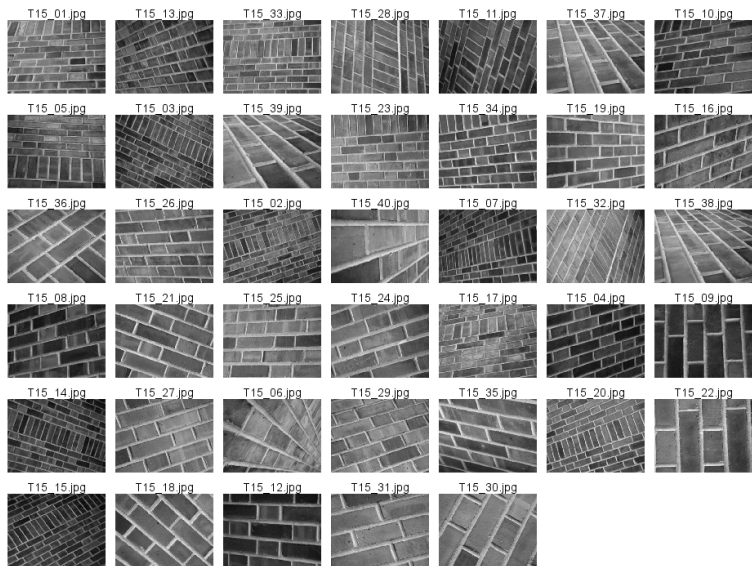
- Invariance à l'échelle :

$$\eta_{pq} = \mu_{pq} / \mu_{00}^{(p+q+2)/2}$$

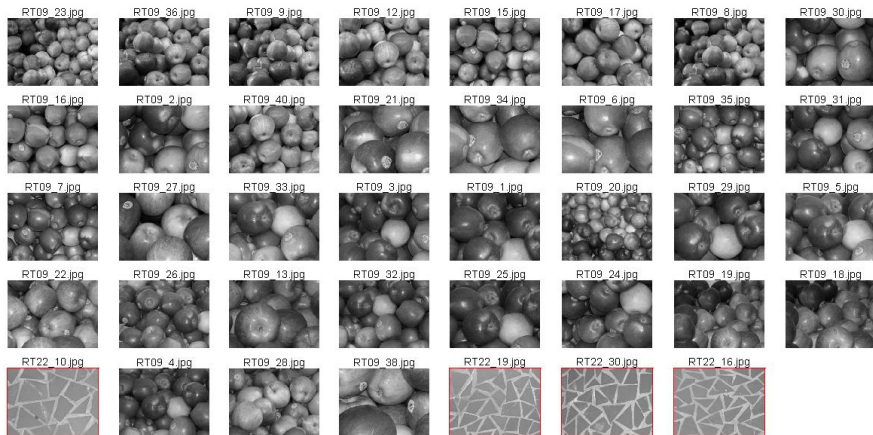
- Invariance aux similitudes : valeurs propres λ_1 et λ_2 de

$$C = \begin{pmatrix} \eta_{20} & \eta_{11} \\ \eta_{11} & \eta_{02} \end{pmatrix}.$$

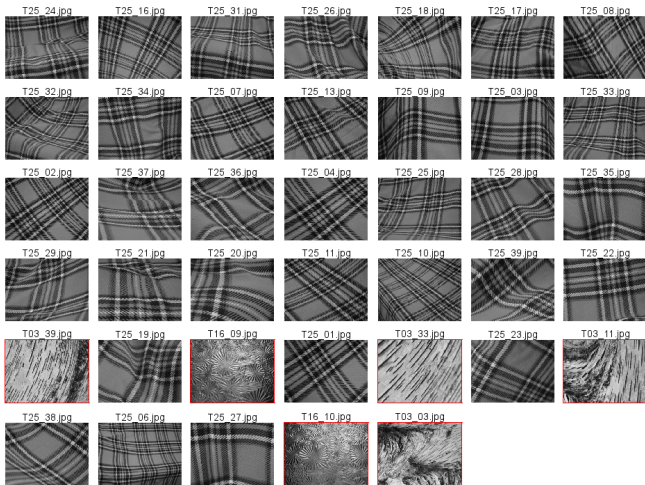
- Invariance affine : $\sqrt{\lambda_1 \lambda_2}$



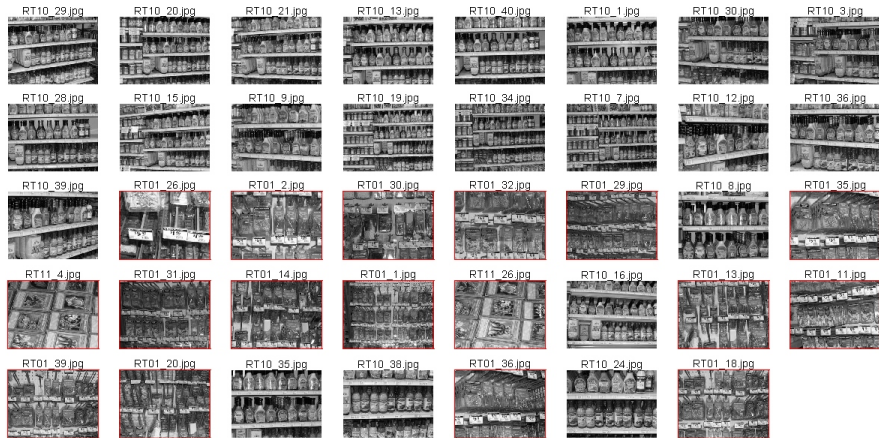
Expérience de recherche par le contenu (25 classes de 40 échantillons)



Expérience de recherche par le contenu



Expérience de recherche par le contenu



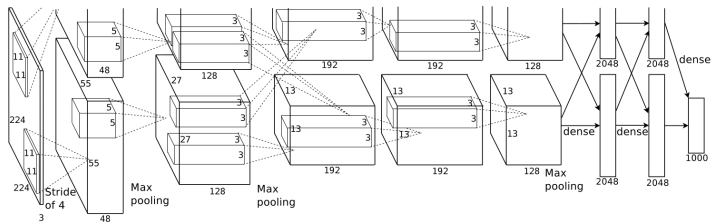
Expérience de recherche par le contenu

Reconnaissance par réseaux de neurones

- Etat de l'art pour l'analyse d'images en un sens large
- Repose sur des réseaux **profonds** et **convolutionnels** (CNNs convolutional neural networks)
- Permet la reconnaissance de textures et de matériaux en environnement non-contrôlés (“in the wild”).
- Nécessite de très grandes quantités d'images **annotées**

Reconnaissance par réseaux de neurones

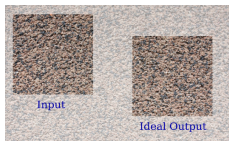
- Idée : **utiliser les couches convolutionnelles d'un CNN comme un ensemble de filtres d'analyse** (Cimpoi-Maji-Vedaldi 2015)
- Repose sur un réseau entraîné sur environ 1M d'images organisées en environ 1000 classes (ImageNet, Krizhevsky et al., 2012)



- Puis apprentissage des textures à partir de 50000 régions, issues de 10000 images, pour 23 classes :
statistiques des réponses aux filtres (par exemple en utilisant des sacs de mots comme précédemment, ou d'autres représentations, e.g. des vecteurs de Fisher)

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



III. B. Galerie



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

Quelques méthodes ad'hoc

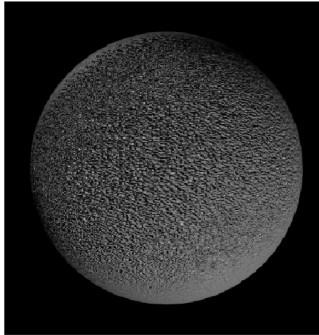
- Stockage d'une image de texture, puis plaquage de la texture avec des conditions aux bords → problèmes topologiques.
- Perturbation des normales (Blinn 1978) :
à partir d'une surface, on calcule en chaque point une direction normale fictive, après perturbation de la surface, puis on utilise cette nouvelle normale dans un modèle de réflexion :

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

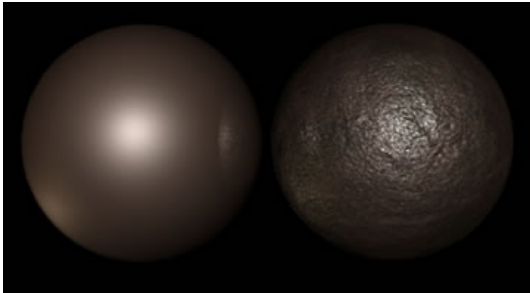
ce qui implique

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

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



Perturbation des normales d'une sphère



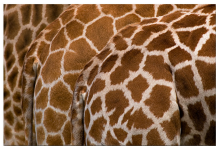
Perturbation des normales d'une sphère
Illustration de <http://www.interstation3d.com>

- 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.

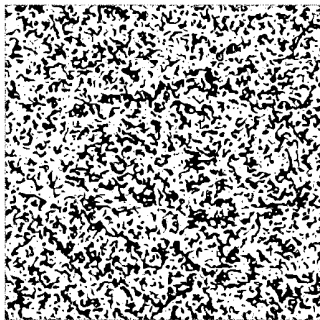
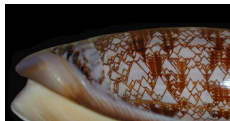
$$\frac{\partial C}{\partial t} = a^2 \Delta C - bC + R.$$

Pour modéliser l'anisotropie, ΔC devient

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



All the same, but different



Reaction-diffusion

Modèles “fractals”

Essentiellement des simulations de généralisations du mouvement brownien : les mouvements browniens fractionnaires (Mandelbrot-Van Ness 1968).

- dimension 1 : V_H processus tel que

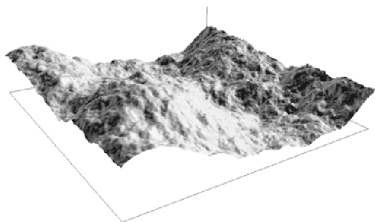
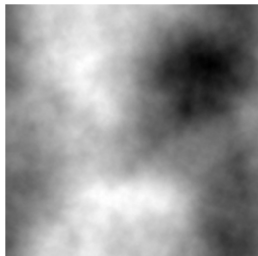
$$V_H(t_1) - V_H(t_2) \sim N(0, |t_2 - t_1|^{2H}).$$

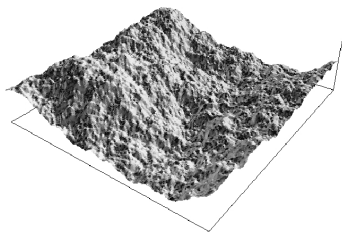
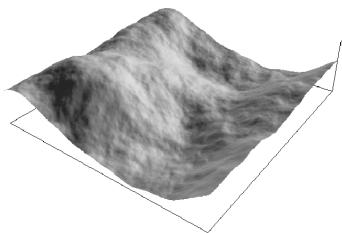
Remarque :

$V_H(t_0) - V_H(t_0 + t)$ 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

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







<http://www.decarpentier.nl>

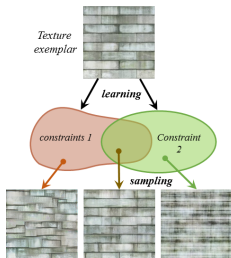
Modèles génériques

Deux grandes familles de méthodes :

- **Synthèse par contraintes statistiques**

Réalisations de modèles aléatoires avec des caractéristiques statistiques inspirées des textures naturelles

→ projections sur des ensembles d'images définis par des contraintes



- **Synthèse par copier-coller de patches**

Images obtenus par ré-échantillonnage de voisinages locaux (patches)

Synthèse par contraintes statistiques

Une digression préliminaire : le modèle “spot noise” [van Wijk 1991]

- On modélise une texture comme l'addition de formes :

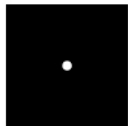
$$f(\mathbf{x}) = \sum a_i h(\mathbf{x} - \mathbf{x}_i),$$

où les a_i sont des v.a. i.i.d, $\{\mathbf{x}_i\}$ un processus de Poisson, h la fonction indicatrice d'une forme.

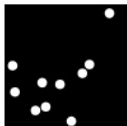
- Converge vers un modèle limite gaussien.
Si on suppose que $E(a_i)=0$, et que le processus $\{\mathbf{x}_i\}$ est de densité ν , alors

$$E|\hat{f}(\omega)|^2 = \nu E(a_i^2) |\hat{h}(\omega)|^2.$$

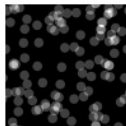
- On synthétise la texture à partir de \hat{h} et de phases aléatoires.
- Source de "bruit" populaire en graphisme [Perlin 1985, Lagae et al. 2009]



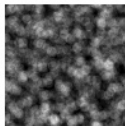
Original
Image



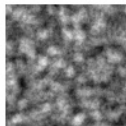
$n = 10$



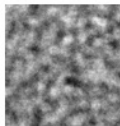
$n = 10^2$



$n = 10^3$



$n = 10^4$

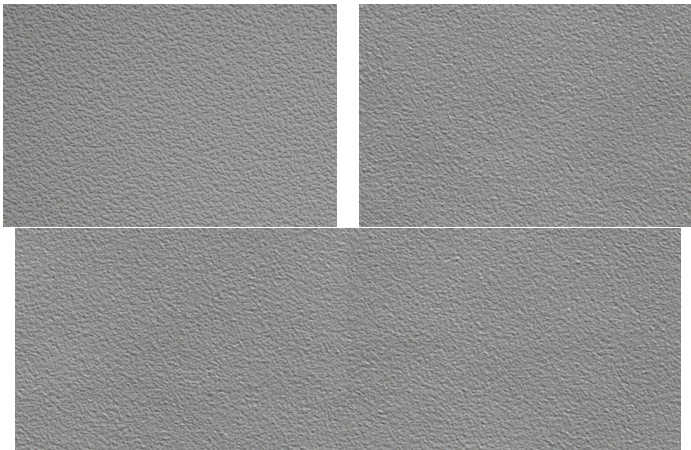


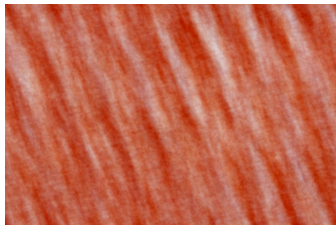
$n = 10^5$

Synthèse par phases aléatoires (Galerne et al. 2011)

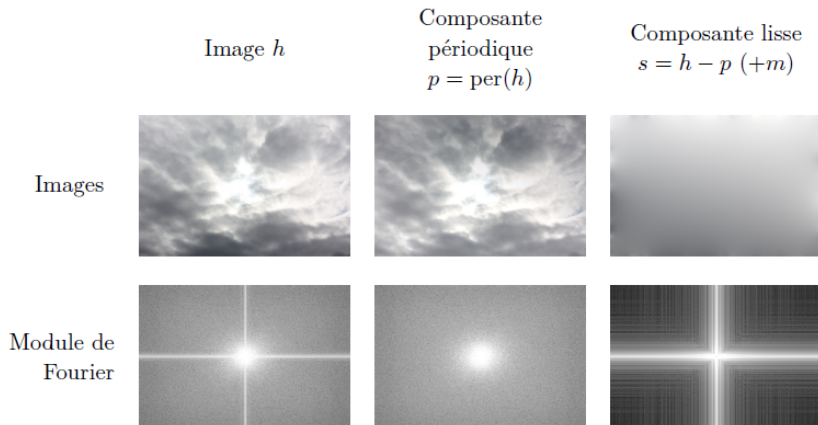
- La forme de base est remplacée par une image couleur
 - Revient à imposer le spectre de la réalisation
 - Deux options pour la synthèse :
 - On impose des phases aléatoires
 - On synthétise le modèle limite gaussien dont on connaît la covariance (donc le spectre)
- résultats visuellement similaires
- On impose une cohérence de phase entre canaux couleurs
 - Suppression des artefacts de non-périodicité (recours à une décomposition en composantes périodiques et régulières)
 - Phases aléatoires : équivalent à la projection orthogonale sur un ensemble d'images à spectre prescrit
 - → Démo en ligne sur *ipol* (*random phase demo* sous google)





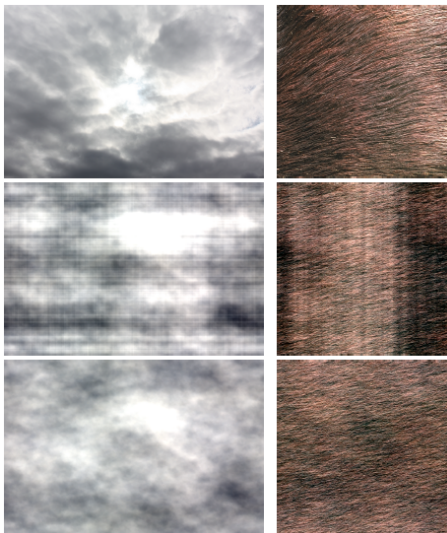


Suppression des composantes non périodiques



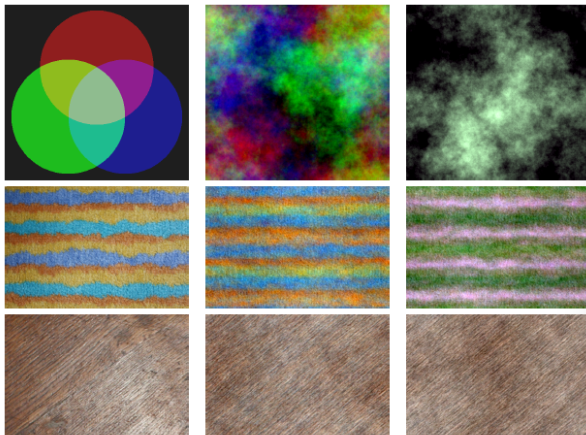
Séparation "periodic + smooth" selon [Moisan 2011]

Suppression des composantes non périodiques



Importance pour la synthèse de la suppression des non stationnarités

Corrélations couleur



Spot h

$$\hat{h} = \begin{pmatrix} |\hat{h}_R| e^{i\varphi_R} \\ |\hat{h}_G| e^{i\varphi_G} \\ |\hat{h}_B| e^{i\varphi_B} \end{pmatrix}$$

BPA

$$\hat{Z} = \begin{pmatrix} |\hat{h}_R| e^{i(\varphi_R + \theta)} \\ |\hat{h}_G| e^{i(\varphi_G + \theta)} \\ |\hat{h}_B| e^{i(\varphi_B + \theta)} \end{pmatrix}$$

“Mauvais” BPA

$$\hat{Z}_M = \begin{pmatrix} |\hat{h}_R| e^{i\theta} \\ |\hat{h}_G| e^{i\theta} \\ |\hat{h}_B| e^{i\theta} \end{pmatrix}$$

On ajoute la même phase aléatoire à chaque canal
→ permet de conserver les corrélations entre canaux

Extensions et applications

- Définition d'un représentant compact (texton)
(Desolneux-Moisan-Ronsin 2012, Galerne et al. 2014)
- Synthèse procédurale rapide (Galerie et al. 2012)
- Modélisation des distributions de température de surface des océans
(Bertrand et al 2014)
- Inpainting de textures gaussiennes (Galerie et al. 2017)
- Initialisation ou contraintes pour modèles plus complexes (Liu et al. 2016, Leclaire et al. 2018)
- Etc.

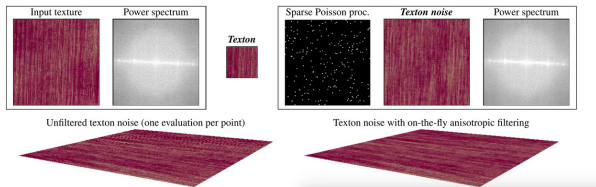
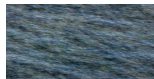
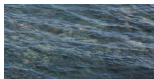
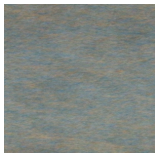
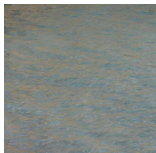
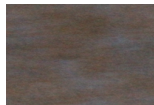
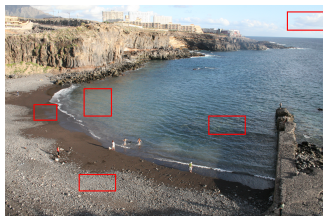


Illustration A. Leclaire



Plusieurs régimes pour une même texture physique



Les champs markoviens

Représentation sous forme de champs de Gibbs avec énergie, inférence des paramètres, échantillonnage de Gibbs.

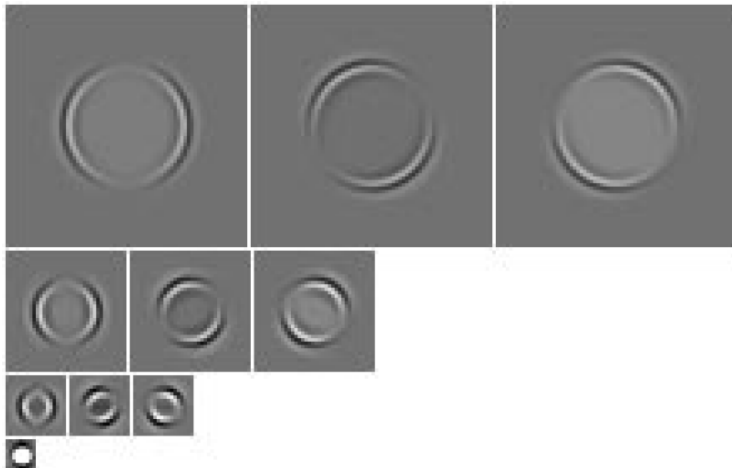
Difficultés : inférence de l'énergie - structures multi-échelles

Méthodes par ondelettes : ajustement d'histogrammes (Heeger-Bergen 1995)


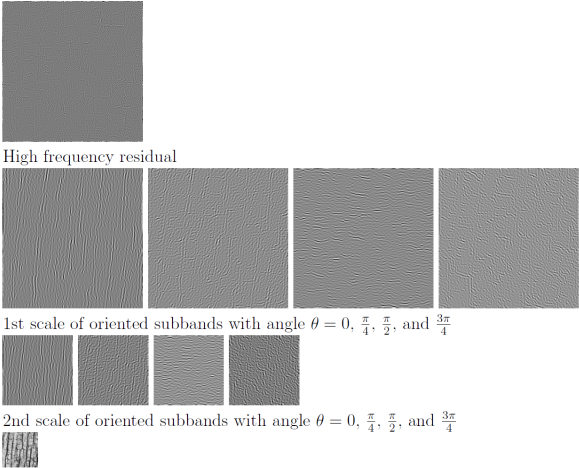
Pour synthétiser une texture B partant d'un échantillon A :

- Calcul d'une décomposition en ondelettes $\{a_{i,j,k}\}$ de A : i orientation, j échelle, k position.
- On part de B bruit blanc, et le décompose en $\{b_{i,j,k}\}$.
- Pour chaque i, j , on modifie l'histogramme de $\{b_{i,j,k}\}$ pour le rendre égal à celui de $\{a_{i,j,k}\}$.
- On reconstruit l'image B .
- On modifie l'histogramme (de couleurs) de B pour le rendre égal à celui de A
- On itère les trois étapes précédentes.

→ Démo en ligne (*heeger ipol* sous google)



Décomposition d'une image de disque dans la base pyramidale "steerable"
(Illustration Briand et al., Ipol)

Original image	Associated steerable pyramid
	 <p data-bbox="385 332 636 356">High frequency residual</p> <p data-bbox="385 567 1023 598">1st scale of oriented subbands with angle $\theta = 0, \frac{\pi}{4}, \frac{\pi}{2},$ and $\frac{3\pi}{4}$</p> <p data-bbox="385 705 1030 736">2nd scale of oriented subbands with angle $\theta = 0, \frac{\pi}{4}, \frac{\pi}{2},$ and $\frac{3\pi}{4}$</p> <p data-bbox="385 795 628 819">Low frequency residual</p>

Décomposition d'une image dans la base pyramidale "steerable"
(Illustration Briand et al., Ipol)

Prescription d'histogramme (cf cours radiométrie)

Pour deux images A et B (à valeurs dans $[0,1]$), on considère le changement de contraste

$$\phi_{BA} : [0, 1] \rightarrow [0, 1],$$

tel que

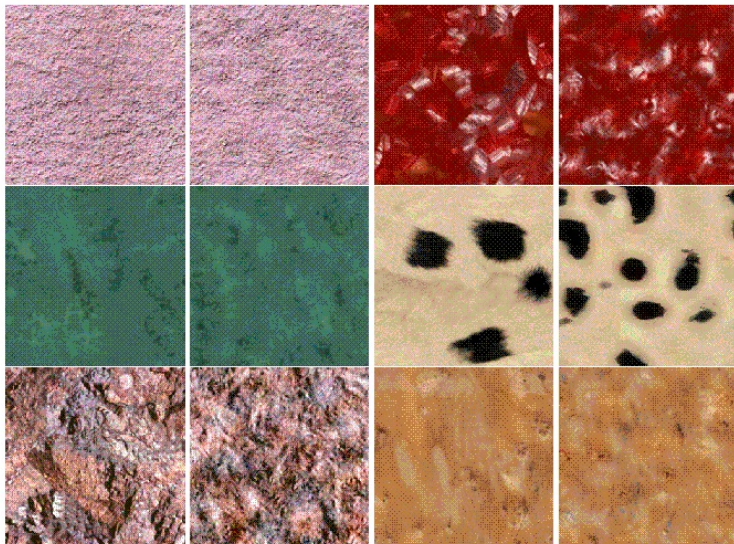
$$h_{\phi_{BA} \circ B} = h_A,$$

où h_I est l'histogramme de I .

Si on note H_I l'histogramme cumulé de I et H_I^{-1} son pseudo-inverse, alors

$$\phi_{BA} = H_A^{-1} \circ H_B$$

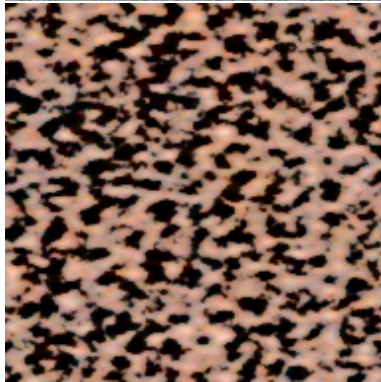
- En pratique, on utilise pas une décomposition sur une base, mais sur une "trame", pour disposer de plus de 3 orientations (pyramide "steerable", voir illustration précédente).
- Cette méthode s'inspire d'observations psychophysiques sur l'importance des reponses à des familles de filtres pour la discrimination des textures,
- Les contours ne sont pas reproduits, car ils correspondent à des relations inter-échelles.



Ajustement des marginales de décompositions en ondelettes. Extrait de Heeger-Bergen 1995



Ajustement des marginales de decompositions en ondelettes. Extrait de Heeger-Bergen 1995



Idem (D mo ipol)

Méthodes par ondelettes : ajustement des statistiques d'ordre supérieur (Simoncelli-Portilla 1998, 2000)

Pour synthétiser une nouvelle image B à partir d'un échantillon A :

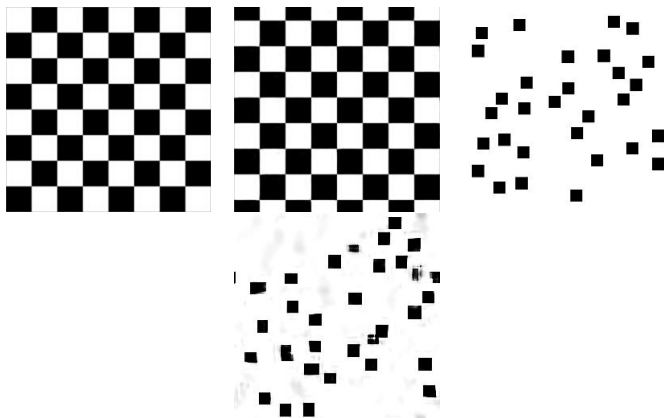
- Calcul d'une décomposition en ondelettes $\{a_{i,j,k}\}$ de A (i : orientation, j : échelle, k : position).
- On part de B égale à un bruit blanc, et le décompose en $\{b_{i,j,k}\}$.
- On projette les $\{b_{i,j,k}\}$ sur des sous-espaces d'éléments ayant les mêmes statistiques (à préciser) que $\{a_{i,j,k}\}$.
- on reconstruit l'image B .

Typiquement, ces statistiques incluent :

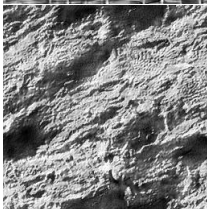
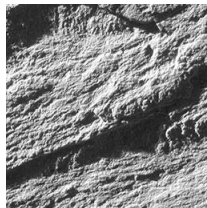
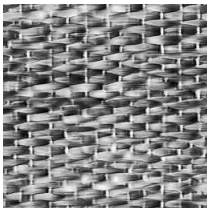
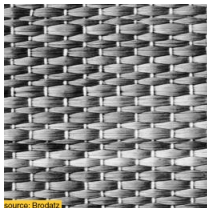
- Moyenne, variance, kurtosis, etc... de B .
- Covariances de $\{|b_{i,j,k}|\}$ et $\{|b_{i,j,k+p}|\}$, pour i, j, p donnés.
- Covariances de $\{|b_{i_1,j_1,k}|\}$ et $\{|b_{i_2,j_2,k}|\}$, pour i_1, i_2, j_1, j_2 donnés.

Ordres de grandeur : quelques orientations (typiquement 4 à 6) ; p de l'ordre de la dizaine, ce qui donne de l'ordre de 1000 paramètres.

L'images B est reconstruite à partir de coefficients $\{|b_{i,j,k}|\}$ obtenues par projection sur des sous-espaces.



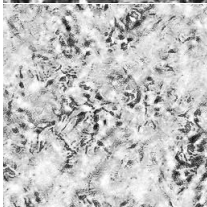
Ajustement de statistiques d'ordres supérieurs de decompositions en ondelettes. Extrait de Portilla-Simoncelli 2000



Ajustement de statistiques d'ordres supérieurs de decompositions en ondelettes. Extrait de Portilla-Simoncelli 2000



Ajustement de statistiques d'ordres supérieurs de decompositions en ondelettes. Extrait de Portilla-Simoncelli 2000



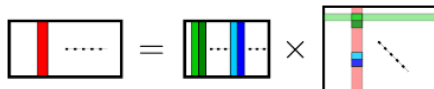
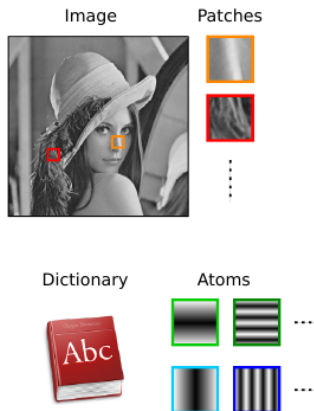
Ajustement de statistiques d'ordres supérieurs de decompositions en ondelettes. Extrait de Portilla-Simoncelli 2000

- Décomposition dans un dictionnaire parcimonieux
- Réseaux convolutionnels profonds

Décompositions parcimonieuses

(Tartavel et al 2015)

- Les ondelettes sont remplacées par un dictionnaire de patches parcimonieux
- Contraintes :
 - Statistiques d'ordre 1 (marginales) sur les coefficients de la décomposition
 - Statistiques d'ordre 1 sur les valeurs de couleur des pixels
 - Spectre de Fourier



$$P = D \times W$$

$$\mathbf{p}_k = \sum_n w_k^n \mathbf{d}^n$$



Illus. G. Tartavel

Dimension reduction

- Less dimension \Rightarrow less coefficients
- Choice of a D by SVD:

$$\min_{D, W} \sum_k \|p_k - Dw_k\|_2^2$$

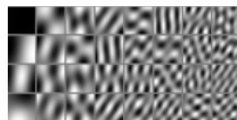
Original image



Approximation 29.5 dB



Dictionary $L = 32$



Illus. G. Tartavel

Learning stage: non-convex, [Elad and Aharon, 2006]

Input image $v \Rightarrow$ dictionary D and weights \tilde{W}

$$\min_{D,W} \|P_v - DW\| + \iota_{\mathcal{C}_s}(W)$$

Sparsity constraints:

$$\mathcal{C}_s = \{W : \|w_k\|_0 \leq S \quad \forall k\}$$

Decomposition stage:

$$\mathcal{D}(u, W) = \frac{1}{Z} \|P_u - DW\| + \iota_{\mathcal{C}_s}(W) + \iota_{\mathcal{C}_f}(W)$$

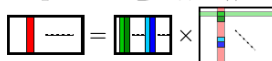
Distribution constraints:

$$\mathcal{C}_f = \{W : \|w^n\|_0 \leq \|\tilde{w}^n\|_0 \quad \forall n\}$$



$$p_k = \sum_n w_k^n d^n$$

$$P = D \times W$$



Cost function:

Function to be minimized:

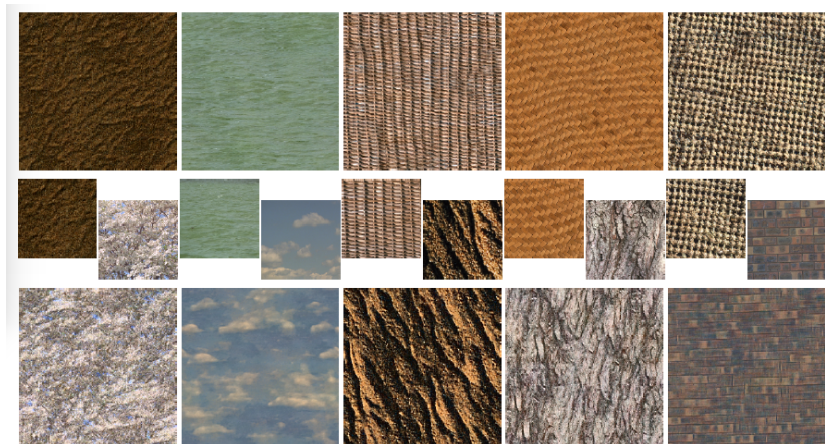
$$E_v(u, W) = \alpha \underbrace{\mathcal{W}^2(\mu_u, \mu_v)}_{\text{Histogram}} + \beta \underbrace{\mathcal{S}^2(u, v)}_{\text{Spectrum}} + \underbrace{\mathcal{D}^2(u, W)}_{\text{Decomposition}}$$

⇒ highly non-convex!

Minimization scheme:

- alternate descent
- w.r.t. u : gradient descent (using u_{hist} and u_{spec})
- w.r.t. W : greedy approximation derived from [Mallat and Zhang, 1993]

Décompositions parcimonieuses

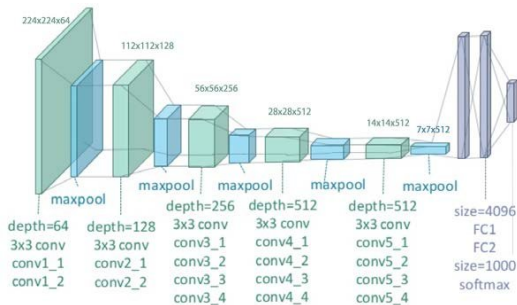


Décompositions parcimonieuses



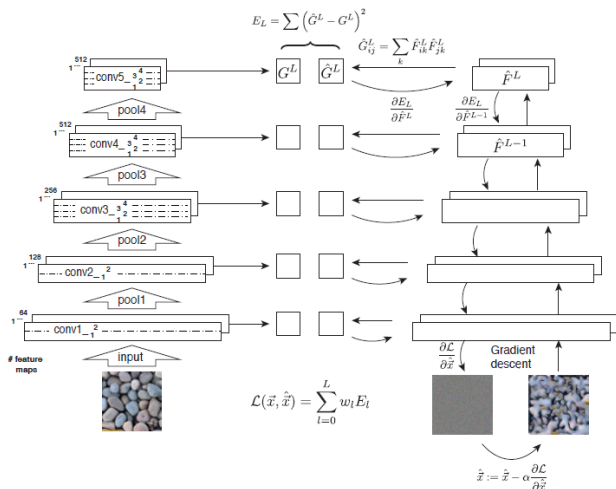
De gauche à droite : exemple, méthode parcimonieuse, Simoncelli-Portilla, phase aléatoire.

- On est toujours dans le cadre de contraintes statistiques
- Les filtres (ondelettes) ou dictionnaires parcimonieux sont remplacés par les réponses à un réseau de neurones
- Les statistiques retenues sont les corrélations entre différents attributs (features) au sein d'une même couche d'un réseau de neurones
- Le réseau utilisé est VGG19 (Simonyan-Zisserman 2014), entraîné pour la classification d'images

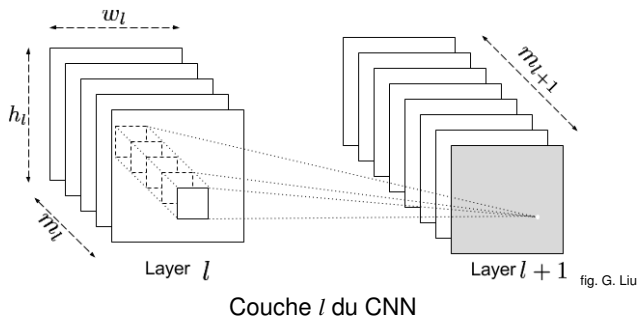


VGG architecture, from C.K. Yang et al. 2018

Réseaux profonds convolutionnels (Gatys-Ecker-Bethge 2015)



Idee : imposer les corrélations entre les différents filtres à la même position



Attributs de texture (features)

Pour une texture exemple I la carte d'attributs (feature map) à la couche l est $f^l \in \mathbb{R}^{m_l \times N_l}$.

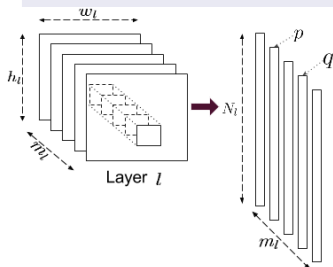
m_l : nombre de filtres, $N_l = w_l \cdot h_l$,

w_l, h_l taille de la carte d'attributs.

Modèle de texture

Inspiré de Simoncelli et al.

On retient les matrices de corrélation $G^l \in \mathbb{R}^{m_l \times m_l}$,



$$\begin{aligned} G_{p,q}^l &= f_p^l \times (f_q^l)^T \\ &= \sum_{i=1}^{N_l} f_p^l(i) \cdot f_q^l(i), \end{aligned} \quad (1)$$

où p, q est l'indice des filtres $p, q \in \{1, \dots, m_l\}$.

Synthèse

On cherche à synthétiser une image ayant les mêmes matrices de corrélation que l'exemple

On part d'un bruit blanc t on définit une fonction de coût à la couche l par

$$E_l = \frac{1}{4(N_l \cdot m_l)^2} \sum_{p=1}^{m_l} \sum_{q=1}^{m_l} \|G_{p,q}^l - \tilde{G}_{p,q}^l\|^2. \quad (2)$$

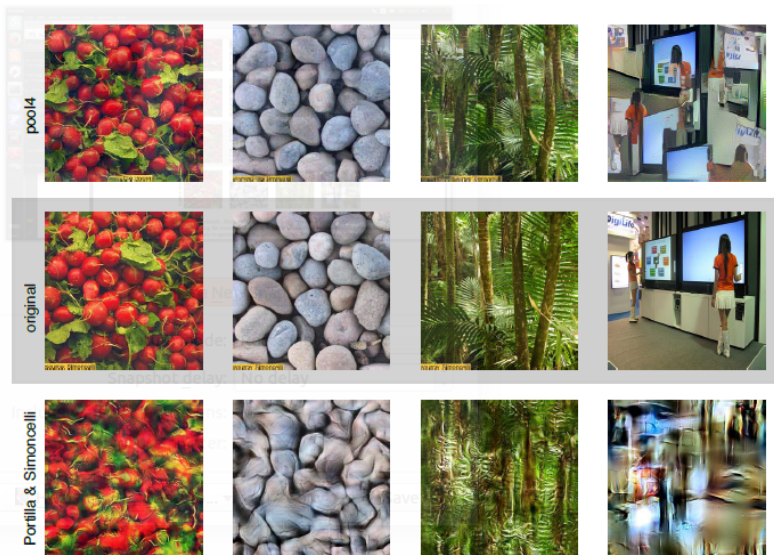
On combine certaines couches

$$\mathcal{L}_{cnn}(I, \tilde{I}) = \sum_{l=0}^L w_l E_l, \quad (3)$$

où w_l sont des poids

Puis on minimise la fonction de coût (back propagation et méthode itérative quasi-Newton, L-BFGS)

Réseaux profonds convolutionnels



Synthesised



Source



Synthesised



Source



Synthesised



Source



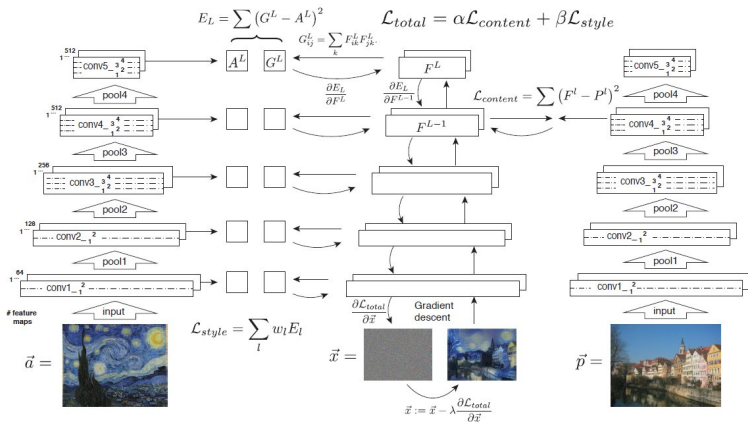


Artefacts de synthèse

Et aussi (à partir d'idées très semblables)



Transfert de style



From Gatys et al. 2015

Transfert de style

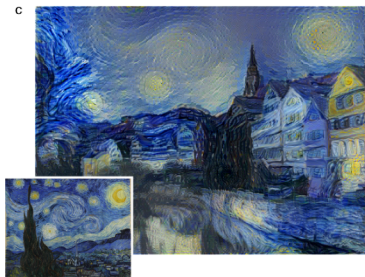


Illustration Gatys et al. 2015

Ne permet pas le respect des dépendances à grande distance



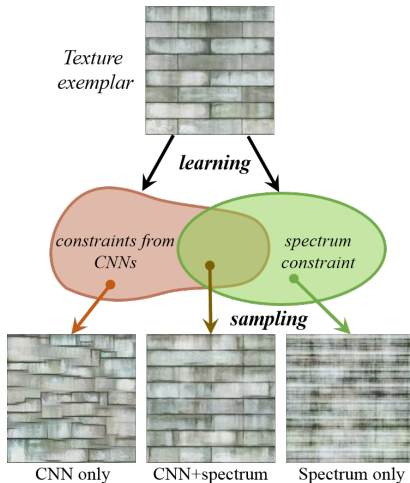


Fig. Gang Liu

FIGURE – Respect simultané des covariances dans le CNN et du spectre (Fourier)
Liu et al. 2016

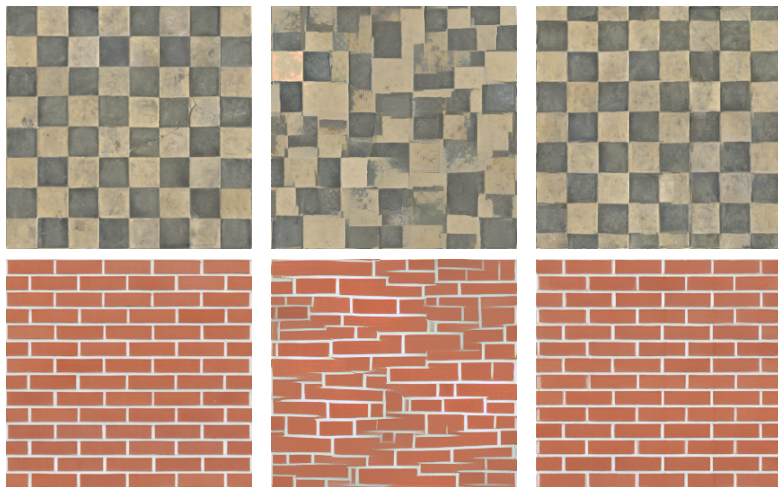


FIGURE – Gauche : original, milieu : CNN, droite : CNN + spectre

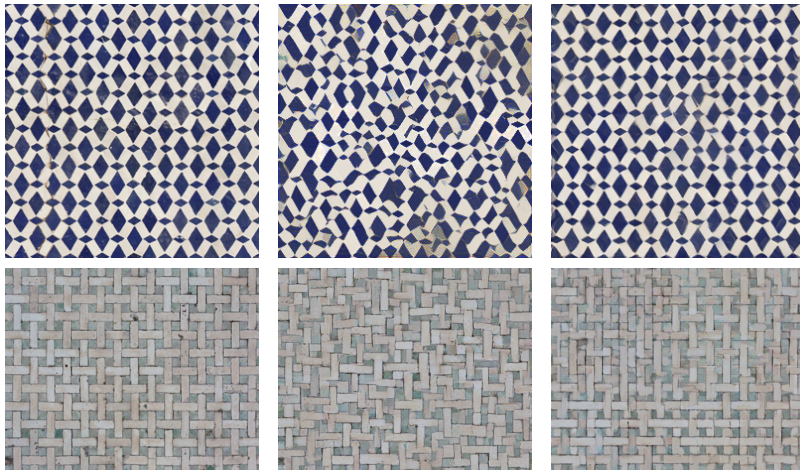


FIGURE – Gauche : original, milieu : CNN, droite : CNN + spectre

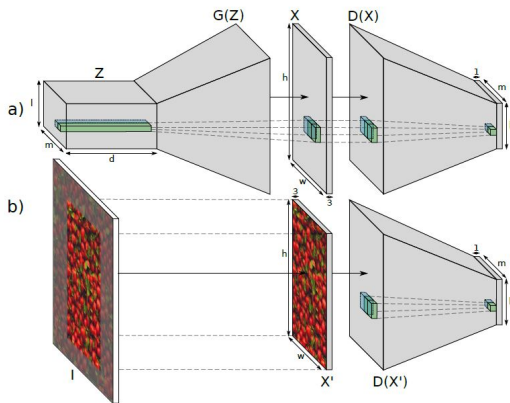
- Alternatives à VGG pré-entraînés : poids aléatoires (He et al 2016, Ustyuzhaninov et al. 2017), architectures alternatives (Wang et al 2019)
- Contraintes statistiques supplémentaires ou alternatives :
 - Spectre (Liu et al 2016, Gonthier et al 2022), covariances intra-features (Sendik et al. 2017)
 - Ordre 1 sur les features (Risser 2017), distributions jointes des features (Heitz et al 2021)
- Représentation multi-échelles (Snelgrove 2017)

Alternative aux approches par optimisation : réseaux génératifs

- Réseaux antagonistes génératifs (Generative Adversarial Networks, GAN) (Jetchev et al. 2016, 2017, Zhou et al. 2018, etc.)
- Entraînement alternatif (Ulyanov et al., 2016, 2017)
- Multi-échelles + GAN : SinGAN (Shaham et al. 2019)
- Auto-encodeur (Chandra et al. 2017)
- Synthèse “universelle” : par normalisation des features (AdaIN, Huang-Belongie 2017), par matching d’histogramme (Risser 2020)
- Synthèse par automates cellulaires (Mortvintsev et al. 2021)
- Modèles de diffusion (texture 3D Chen et al 2023, Cherel et al 2023)
- Etc.

Synthèse par GAN

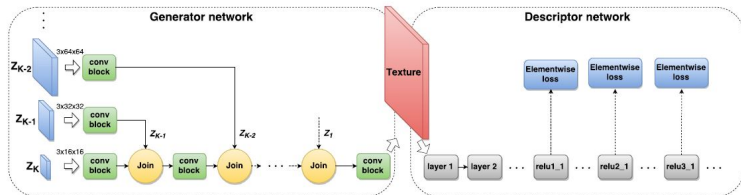
Jetchev et al 2017



Extrait de Jetchev et al.

Approche générative multi-échelles

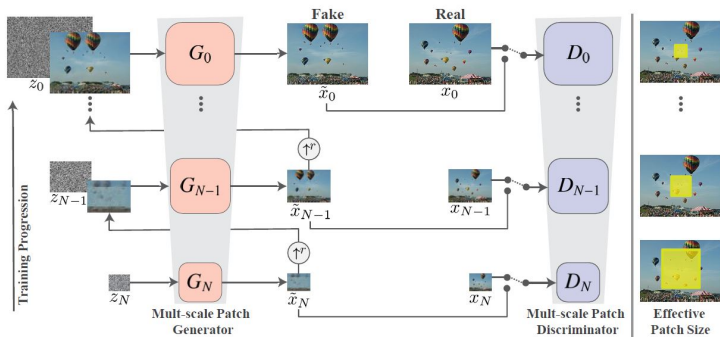
Ulyanov et al 2016,2017



Extrait de

Ulyanov et al.

SinGAN (Shaham et al. 2019)

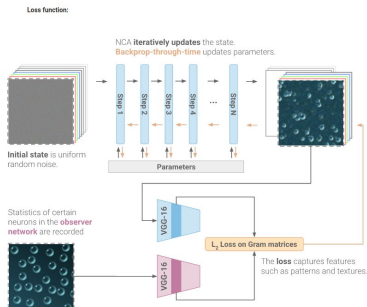
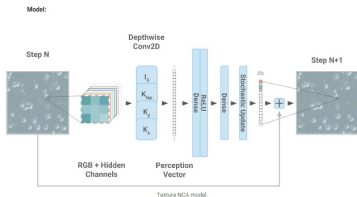


Extrait de Shaham et al.

Synthèse par automates cellulaires

Mordvintsev et al., 2021

- Réseau de neurones peu profond dont les filtres sont des opérateurs différentiels simples (gradient et laplacien)
- Les poids sont appris pour chaque exemple (pour minimiser une loss liée aux matrices de Gram VGG)
- Une texture est synthétisée en faisant passer un bruit blanc dans le réseau
- On itère



Extrait de Mordvintsev et al.

- Méthodes précédentes par contraintes : reposent sur des **heuristiques pour l'échantillonnage** (e.g. projection d'un bruit blanc)
- Certaines méthodes cherchent à optimiser la diversité des synthèses en échantillonnant selon une distribution **maximisant l'entropie** :
FRAME Zhu et al. 1998, de Bortoli et al. 2019
- Echantillonnage selon une mesure de Gibbs permettant de respecter les contraintes en moyenne (modèles macro-canoniques)

Synthèse par copier-coller de patches

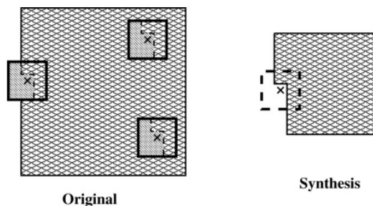
Synthèse par copier-coller de patches

- On considère l'ensemble des patches d'une image exemple : D ,
- On échantillonne les éléments de D pour synthétiser une nouvelle image de texture
- La synthèse s'effectue généralement de manière gloutonne (greedy)

Première approche (Efros-Leung 1999)

Idée :

- considérer la texture comme la réalisation d'un **champs de Markov**
- au lieu d'échantillonner à partir d'un modèle paramétrique (dont les paramètres sont appris sur la texture), on **échantillonne directement sur une image exemple A**
- la synthèse s'effectue pixel par pixel, de manière gloutonne
- 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. 1. *The nonparametric resampling algorithm.*

(fig : Levina-Bickel 2006)

- A partir d'un exemple A , on cherche à synthétiser B
- On se donne une imagerie carrée $F = \{a_1, \dots, a_K\}$ (un patch).
- Pour p un pixel de A , soit $N_A(p) = \{A(p + a_1), \dots, A(p + a_K)\}$, le patch placé en p
- On définit une distance entre patches, 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 synthétiser la valeur en un pixel p' de B on considère un ensemble de **pixels candidats dans A**

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

pour $\epsilon > 0$.

- on tire $p \in O_\epsilon$ et on affecte la valeur $B(p') = A(p)$

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.

→ Démo en ligne (*efros ipol* sous google)

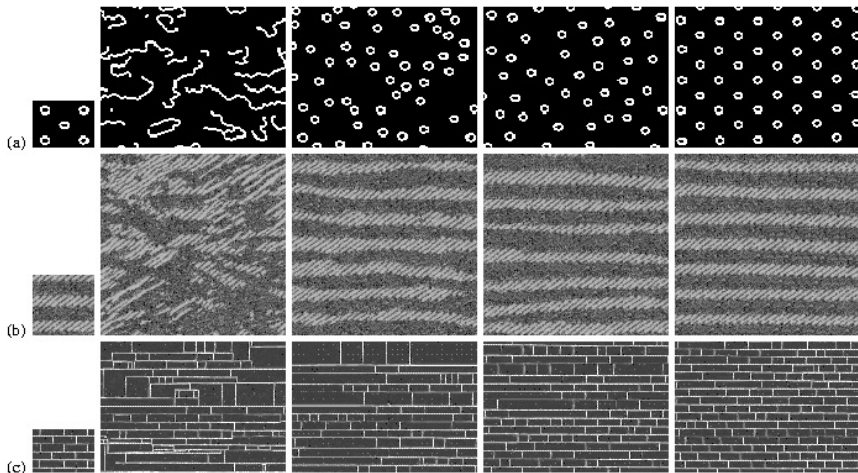
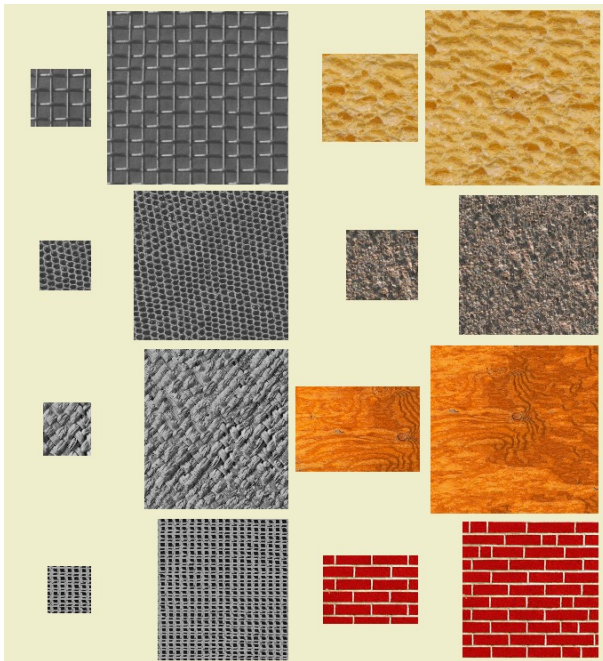


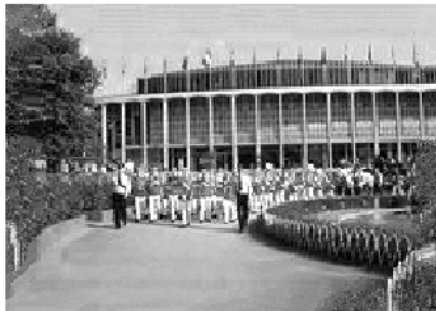
Figure 2. Results: given a sample image (left), the algorithm synthesized four new images with neighborhood windows of width 5, 11, 15, and 23 pixels respectively. Notice how perceptually intuitively the window size corresponds to the degree of randomness in the resulting textures. Input images are: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



ut it becomes harder to lau
ound itself, at "this daily
ving rooms," as House Der
scribed it last fall. He fai
at he left a ringing questi
ore years of Monica Lewir
inda Tripp?" That now see
Political comedian Al Frat
xt phase of the story will

HE THUS THE TOTAL COULD BE 11, at this as Lew at be y
at nda trears coune Tring rooms," as Heft he fast nd it l
ars dat noears outseas ribed it last nt. hest bedian Al. F
eonical Horn d it h Al. Heft ars of, as da Lewindail f l
lian Al Ths," as Lewing questies last aticarsticall. He
is dian Al last fal counda Lew, at "this dailyears dily
edianicall. Hooxewing rooms," as House De fale f De
und itical counoestscribed it last fall. He fall. Hefft
rs oroheoned it nd it he left a ringing questica Lewin.
icars coecoms," astore years of Monica Lewinow seee
a Thas Fring roome stooniscat nowea re left a roouse
bouestof MHe lelf a Lést fast ngine laumesticars Hef
ud it rip?" THousef, a ringind itsonestid. it a ring que:
astical cois ore years of Mounq fall. He ribof Mouse
ore years ofanda Tripp?" That hedian Al Lest fasee yea
nda Tripp? Political comedian Alét he few se ring que
olitical cona re years of the storears ofàs l Frat nica L
res Lew se lest a rime l He fas quest nging of, at beou

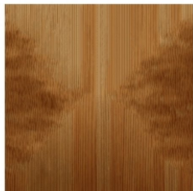
Extrait de Efron-Leung 1999



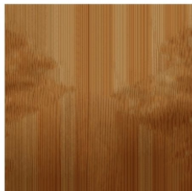
Extrait de Efros-Leung 1999



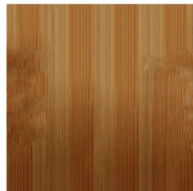
(a) Input
image



(b) $w = 9$



(c) $w = 15$



(d) $w = 21$

Input Image



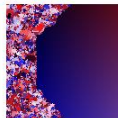
Output Image

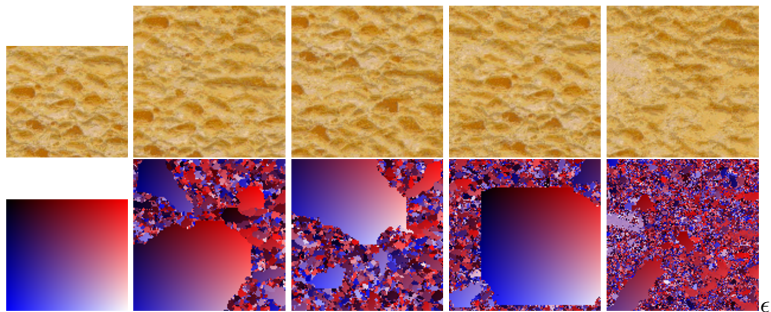


Position Map

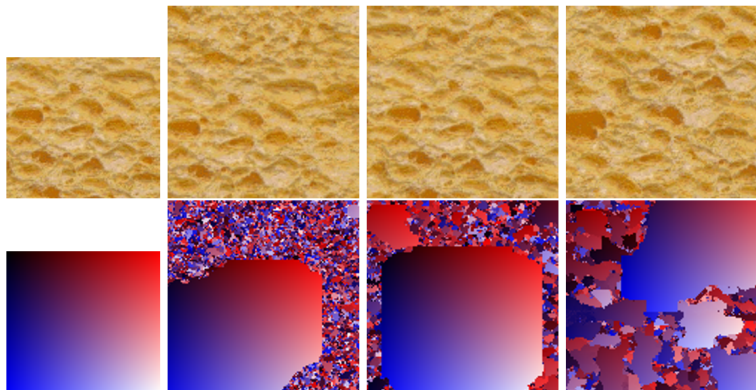


Copy Map





0,05 ; 0,1 ; 0,3 ; 0,5



Patch size : 5, 7, 9

De nombreuses variantes ont été proposées pour la synthèse de texture non-paramétrique par *patches*

- **Quantification vectorielle**
Fast Texture Synthesis using Tree-structured Vector Quantization L. L. Wei et M. Levoy, 2000
- **Synthèse** directement **par patches** (au lieu de pixels)
Image quilting for texture synthesis and transfer, A. Efros and B. Freeman, 2001
→ Démo IPOL
- **Implémentations parallèles** ; GPU
State of the art in example-based texture synthesis, LY Wei, S Lefebvre, V Kwatra, G Turk, 2009
- **Génération des patches** par modélisation par mélange de gaussienne (Raad et al. 2016)
- **Distance de transport sur les patches**
A texture synthesis model based on semi-discrete optimal transport in patch space B. Galerne, A. Leclaire, J. Rabin, 2018

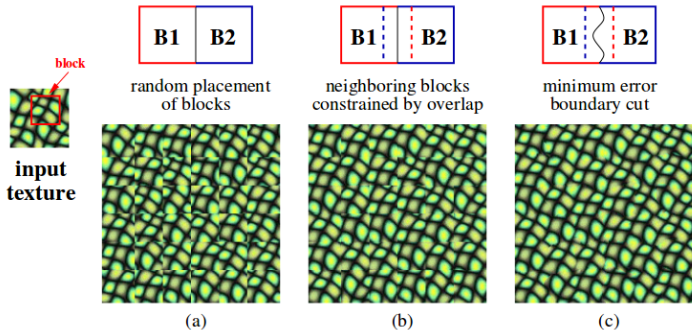
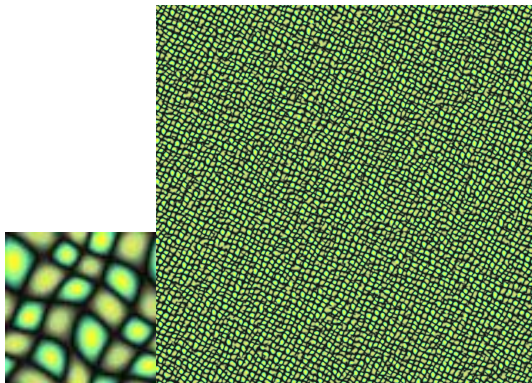


Illustration de Eros-Freeman 2001

Les patches sont accolés séquentiellement avec une procédure de recollement (frontière entre patches minimisant les discontinuités)

- Initialisation : champs gaussien
- Puis : perturbations pour imposer aux patches de l'images de suivre la distribution apprise sur l'exemple
- Résolution d'un problème de transport optimal semi-discret sur les patches
- En multi-échelles



Suite : cours sur l'inpainting