

CSC_4MI04 – Cours n°1

Introduction & Modèles

ENSTA 2^e année
Mineures Info/IAC et Méca/SI

Antoine MANZANERA
ENSTA-Paris / U2IS



ENSTA



CSC_4MI04 – Objectifs du cours

Ce cours, commun aux mineures Info / IAC et Méca / SI, présente les concepts essentiels à l'analyse de bas niveau en vision par ordinateur: les outils fondamentaux de l'analyse, l'échantillonnage et le codage, la représentation des images et des vidéos, l'extraction de caractéristiques.

Il présente aussi les principales techniques utilisées pour traiter et analyser les images avant leur interprétation en termes de sémantique ou de reconstruction 3d, qui seront utiles pour les cours orientés vers la perception des systèmes autonomes en 3e année : parcours IA et ROB (Info) et MID (Méca).

Un tiers du cours est consacré aux méthodes de classification / apprentissage dédiées aux données images, dans un cadre supervisé (bayésien, k-ppv, réseaux convolutionnels) et non supervisé (ACP, K-moyennes, auto-encodeurs).

CSC_4MI04 – Programme 1/2

CRÉNEAU	DESCRIPTION DE LA SÉANCE	PROFESSEUR	RESSOURCES
Vendredi 7/2/25 9:00-12:15	<p>Introduction : Modélisation et représentation des images</p> <ul style="list-style-type: none"> • Échantillonnage et Quantification • Outils fondamentaux (Convolutions, Histogrammes, Transformées de Fourier)... • TD / TP: Convolution et Filtrage 	Antoine MANZANERA (Cours+TP)	<ul style="list-style-type: none"> • Polycopié Chap 1 : "Les images numériques" • Polycopié Chap 2 : "Filtrage et Restauration" • Transparents du cours "Introduction et Modèles"
Vendredi 14/2/25 9:00-12:15	<p>Détection et Représentation de Caractéristiques (partie 1)</p> <ul style="list-style-type: none"> • Cours (9h-11h) <ul style="list-style-type: none"> ◦ Géométrie différentielle pour les images, Dérivées multi-échelles, ◦ Contours, Points d'intérêt, Segmentation... ◦ Descripteurs et Indexation (1/2) • TP (11h15-12h15) <ul style="list-style-type: none"> ◦ Filtrage et Extraction de Caractéristiques (1/2) 	Antoine MANZANERA (Cours+TP) Marwane HARIAT (TP)	<ul style="list-style-type: none"> • Polycopié Chap 3 : "Caractéristiques multi-échelles" • Transparents du cours "Caractéristiques"
Vendredi 21/2/25 9:00-12:15	<p>Détection et Représentation de Caractéristiques (partie 2)</p> <ul style="list-style-type: none"> • Cours (9h-10h) <ul style="list-style-type: none"> ◦ Descripteurs et Indexation (2/2) ◦ Principes de l'indexation d'images et de la recherche par le contenu... • TP (10h15-12h15) <ul style="list-style-type: none"> ◦ Filtrage et Extraction de Caractéristiques (2/2) 	Antoine MANZANERA (Cours+TP) Marwane HARIAT (TP)	<ul style="list-style-type: none"> • Enoncé du TP "Features" • Logiciel pour le TP
Vendredi 7/3/25 9:00-12:15	<p>Classification et Apprentissage de données images - partie 1</p> <ul style="list-style-type: none"> • Cours (9h-10h30) <ul style="list-style-type: none"> ◦ Réduction de dimension et Clustering ◦ Apprentissage Bayésien ◦ k-NN, SVM, Random Forest • TP (10h45-12h15) <ul style="list-style-type: none"> ◦ Classification de pixels supervisée (Modèles bayésiens) et non supervisée (K-Means) 	Antoine MANZANERA (Cours+TP) Marwane HARIAT (TP)	<ul style="list-style-type: none"> • Transparents du cours "Classification et Apprentissage (voir Partie 1)" • Enoncé du TP2 • Logiciel pour le TP2

CSC_4MI04 – Programme 2/2

Vendredi 14/3/25 8:30--11:45	<p>Classification et Apprentissage de données images - partie 2</p> <ul style="list-style-type: none">• Cours (8:30-10h)<ul style="list-style-type: none">◦ Réseaux de neurones (1/2)• TP (10h15-11h45)<ul style="list-style-type: none">◦ Réseaux convolutionnels : Classification d'images par CNN (1/2)	Antoine MANZANERA (Cours+TP) Marwane HARIAT (TP)	<ul style="list-style-type: none">• Transparents du cours "Classification et Apprentissage (voir Partie 2)"
Vendredi 21/3/25 9:00--12:15	<p>Représentation et Détection d'objets</p> <ul style="list-style-type: none">• Représentations d'objets visuels, Sacs de mots visuels, Fenêtre glissante.• Transformées de Hough pour la reconnaissance de formes paramétrées. <p>Vidéo : analyse du mouvement :</p> <ul style="list-style-type: none">• Intérêt du flot optique, Techniques différentielles et fréquentielles,• Calcul du Flux optique épars (Lucas et Kanade) et dense (Horn et Schunck).• Codage et compression Mpeg de vidéos.	Antoine MANZANERA	<ul style="list-style-type: none">• Polycopié Chap 4 : "Formes paramétrées"• Transparents du cours "Modèles pour la Détection d'Objets"• Transparents du cours "Estimation du mouvement"• Vidéos du cours "Mouvement" : Guepardo et Anafi
Vendredi 28/3/25 9:00--12:15	<p>Classification et Apprentissage de données images - partie 3</p> <ul style="list-style-type: none">• Cours (8h30-10h)<ul style="list-style-type: none">◦ Réseaux de neurones (2/2)• TP (10h15-11h45)<ul style="list-style-type: none">◦ Réseaux convolutionnels : Classification d'images par CNN (2/2)	Antoine MANZANERA (Cours+TP) Marwane HARIAT (TP)	<ul style="list-style-type: none">• Transparents du cours "Classification et Apprentissage (voir Partie 2)"• Enoncé du TP "CNN"• Notebook Jupyter pour le TP

→ <https://perso.ensta-paris.fr/~manzaner/Cours/MI204/>

Cours n°1 : Introduction

I Traitement d'Images et domaines connexes

- I-1 Historique du Traitement d'images (TI)
- I-2 TI et vision par ordinateur
- I-3 TI et Intelligence Artificielle
- I-4 TI et perception visuelle

II Introduction aux images numériques

- II-1 Modalités
- II-2 Vocabulaire
- II-3 Échantillonnage et quantification

III Les modèles formels du TI

- III-1 Le modèle linéaire : la convolution...
- III-2 Le modèle fréquentiel : la transformée de Fourier, l'échantillonnage, ondelettes...
- III-3 Le modèle statistique : l'histogramme, la quantification, l'entropie,...
- III-4 Le modèle différentiel : gradients, isophotes, équations différentielles,...
- III-5 Le modèle ensembliste : morphologie mathématique,...
- III-6 Le modèle discret : maillage, connexité, distances,...

Applications du TI - 1

ROBOTIQUE

PRAXITÈLE (INRIA)



conduite automatique

- ✗ évitement d'obstacle
- ✗ reconstruction 3D
- ✗ environnement non structuré / hostile
- ✗ temps réel primordial

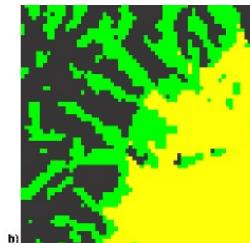
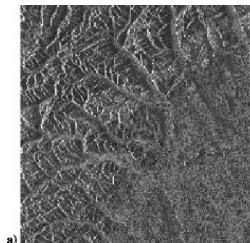
AIBO (SONY)



suivi et préhension
d'objets en robotique
mobile

TELEDETECTION

Extraction automatique du relief à partir d'images satellite radar à synthèse d'ouverture (Univ. Bonn)



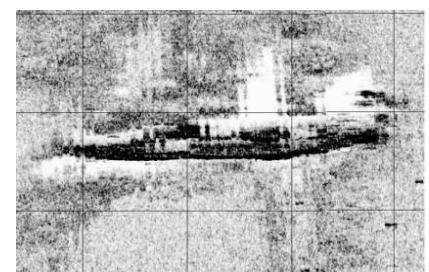
environnement

Image d'un carrefour
(Univ. Karlsruhe)



surveillance du
trafic routier

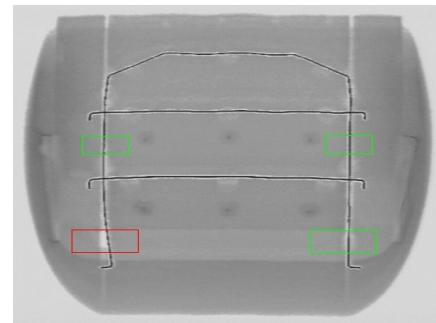
Image sonar d'une épave de sous-marin
reposant à 5000m de profondeur (Nauticos)



exploration des fonds marins

CONTÔLE INDUSTRIEL

vérification structure métallique des sièges
autos par radiographie X (TEKCIM)



détection de défaut

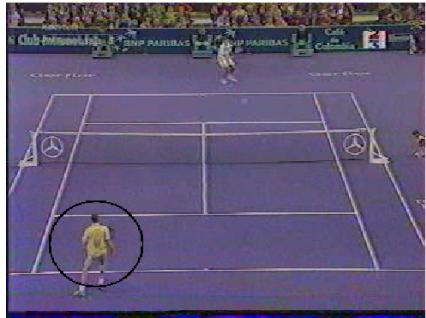
- ✗ environnement maîtrisé

- ✗ gros volume de données
- ✗ temps différé possible

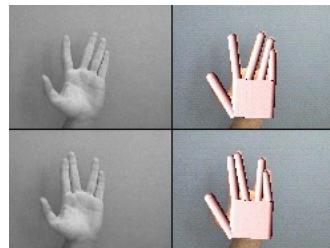
Applications du TI - 2

MULTIMEDIA

Reconnaissance d'action dans un match de tennis pour l'indexation vidéo (INRIA)



Reconnaissance et resynthèse de geste pour le codage et la transmission du langage signé (INT Évry)



codage et compression

recherche d'images dans une base de données à partir d'une requête visuelle (ENSEA Cergy-Pontoise)



indexation d'images

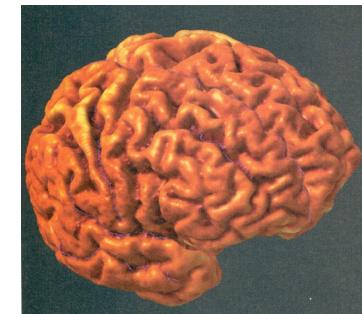
restauration d'images



Figure 42. Portrait of a Lady, painted in 1470 by Petrus Christus:
a) present appearance, showing cracks and fading of the colors;
b) application of a median filter to fill in the dark cracks, and color compensation to adjust the colors.

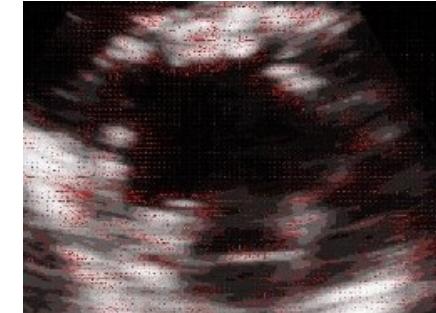
MÉDECINE

modèle 3D de cerveau reconstitué à partir de coupes d'images acquises par résonance magnétique (Univ. Québec)



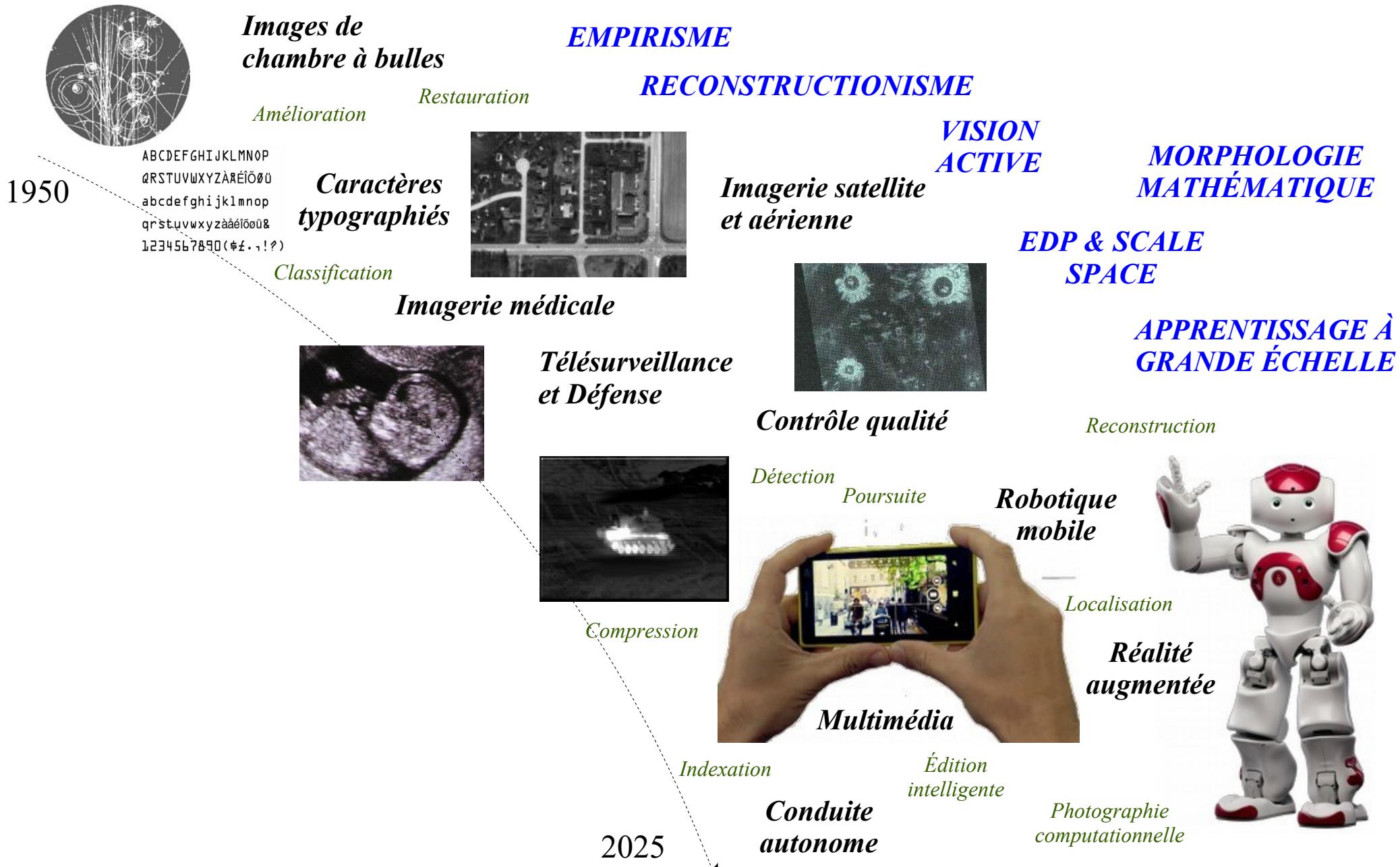
téléchirurgie et ChAO

estimation de mouvement sur séquence d'images échocardiographique (INRIA)

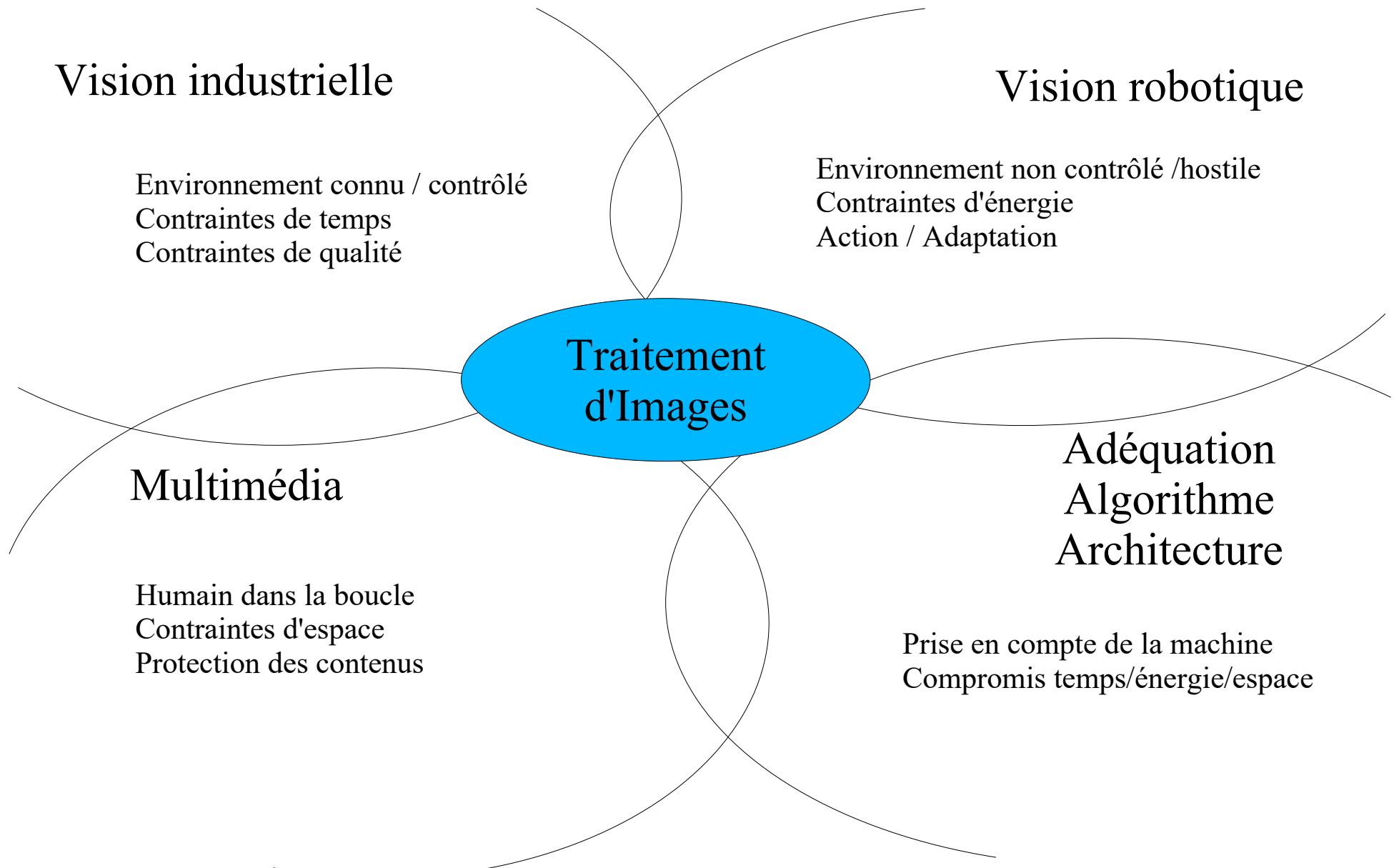


Aide au diagnostic

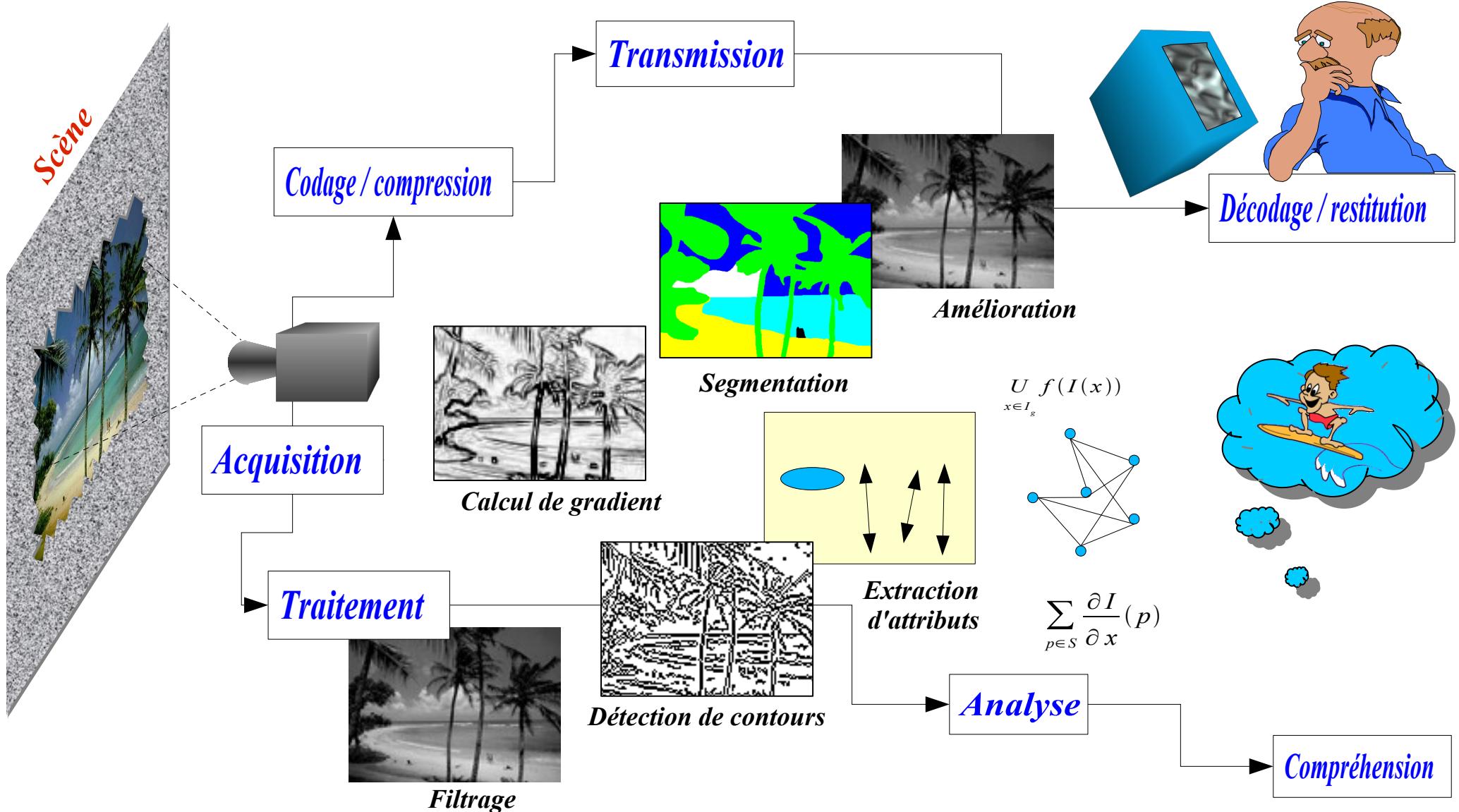
Une (très brève) histoire du TI...



TI & vision par ordinateur



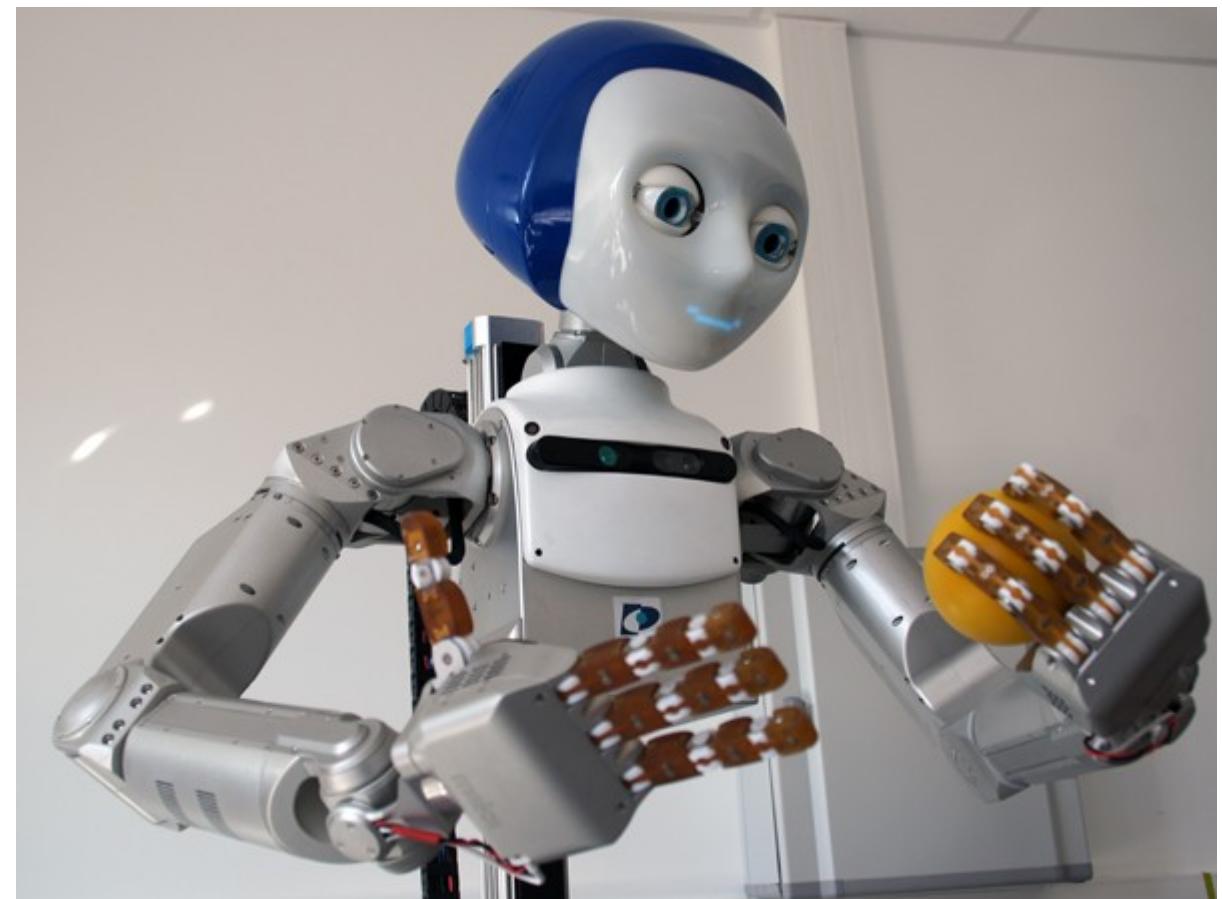
Systèmes à base de TI



Vision pour l'IA...

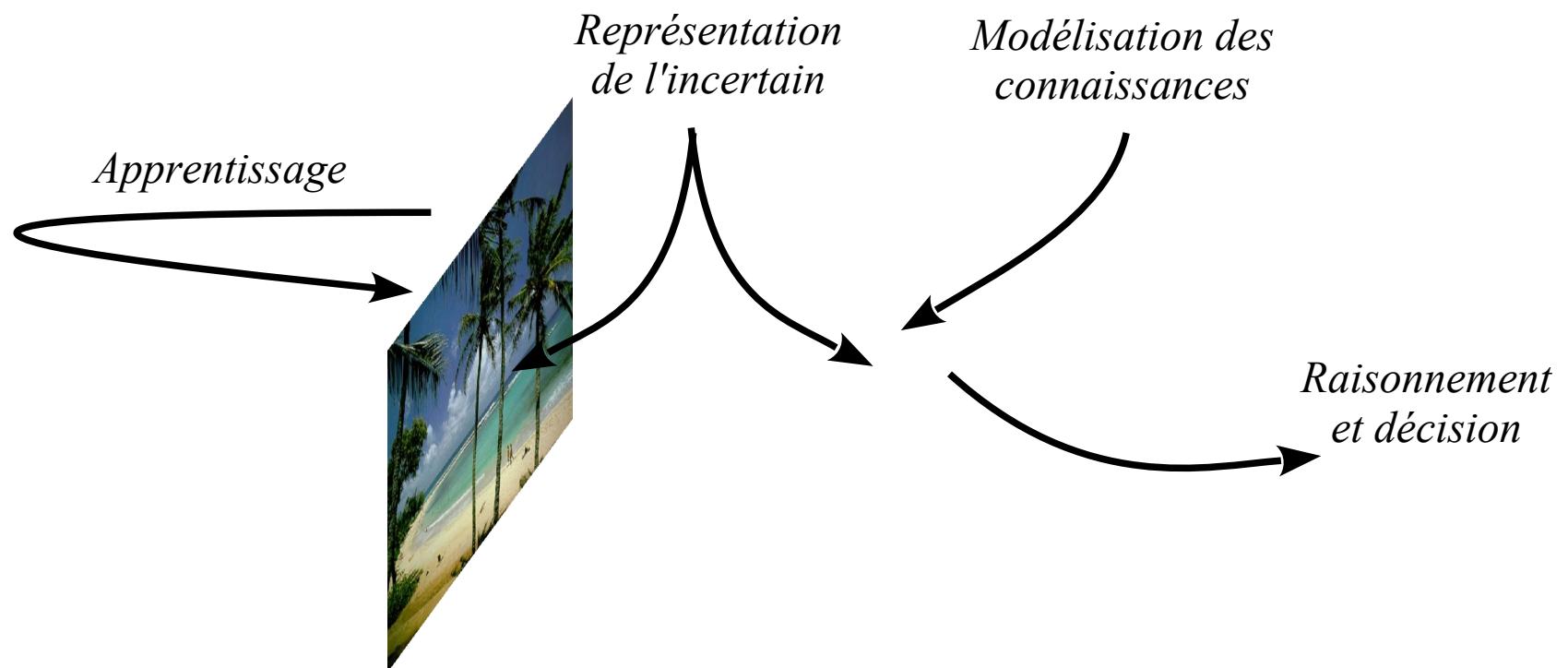
Dans la conception moderne de l'Intelligence Artificielle dite *située* (i.e. mise en situation) ou *incarnée* (i.e. introduite dans un « acteur »), la machine *agit* sur le monde extérieur, éventuellement *se déplace*, et aussi *perçoit* son environnement pour pouvoir s'y *adapter*.

La *vision* est une source extrêmement *riche* d'information, qui permet à la machine de se localiser, reconnaître des objets ou des personnes, à un *coût* faible, une *énergie* raisonnable, et de manière *passive* (i.e. sans émettre de signal).



...et IA pour la Vision !

Réciprocement, le traitement d'images et la vision tirent parti des connaissances et des techniques d'intelligence artificielle pour gérer l'adaptation à un environnement changeant, l'information incertaine, les systèmes hétérogènes de connaissances et les différents niveaux de prise de décision.



TI & perception visuelle

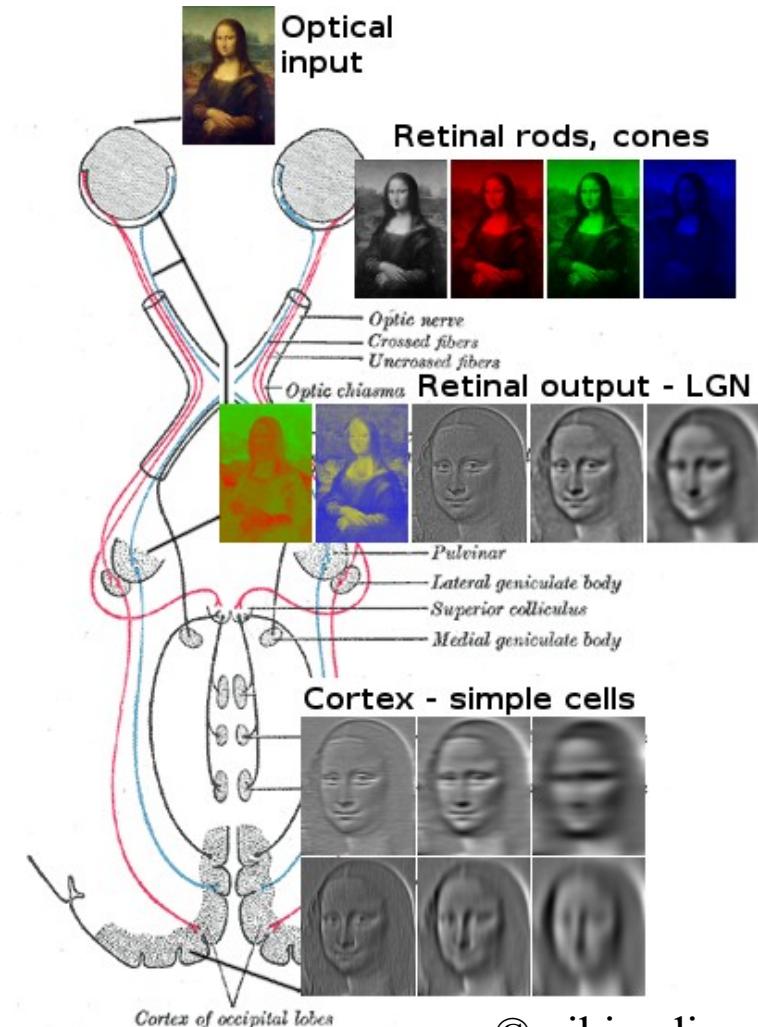
Une difficulté fondamentale de la vision artificielle vient du manque de connaissance profonde des mécanismes qui régissent la compréhension des images dans la nature.

La vision humaine est en effet extrêmement performante (déplacement, lecture, reconnaissance), mais nous n'avons aucun retour conscient sur les mécanismes mis en jeu (à la différence du jeu d'échec par exemple).

En cela l'étude des mécanismes physiologiques et psychologiques de la vision sont une source très importante d'information, et d'inspiration.

Exemples :

- ➔ Traitements rétiniens / traitements corticaux.
- ➔ Mécanisme d'accentuation des contrastes.
- ➔ Multi-résolution et rétine...



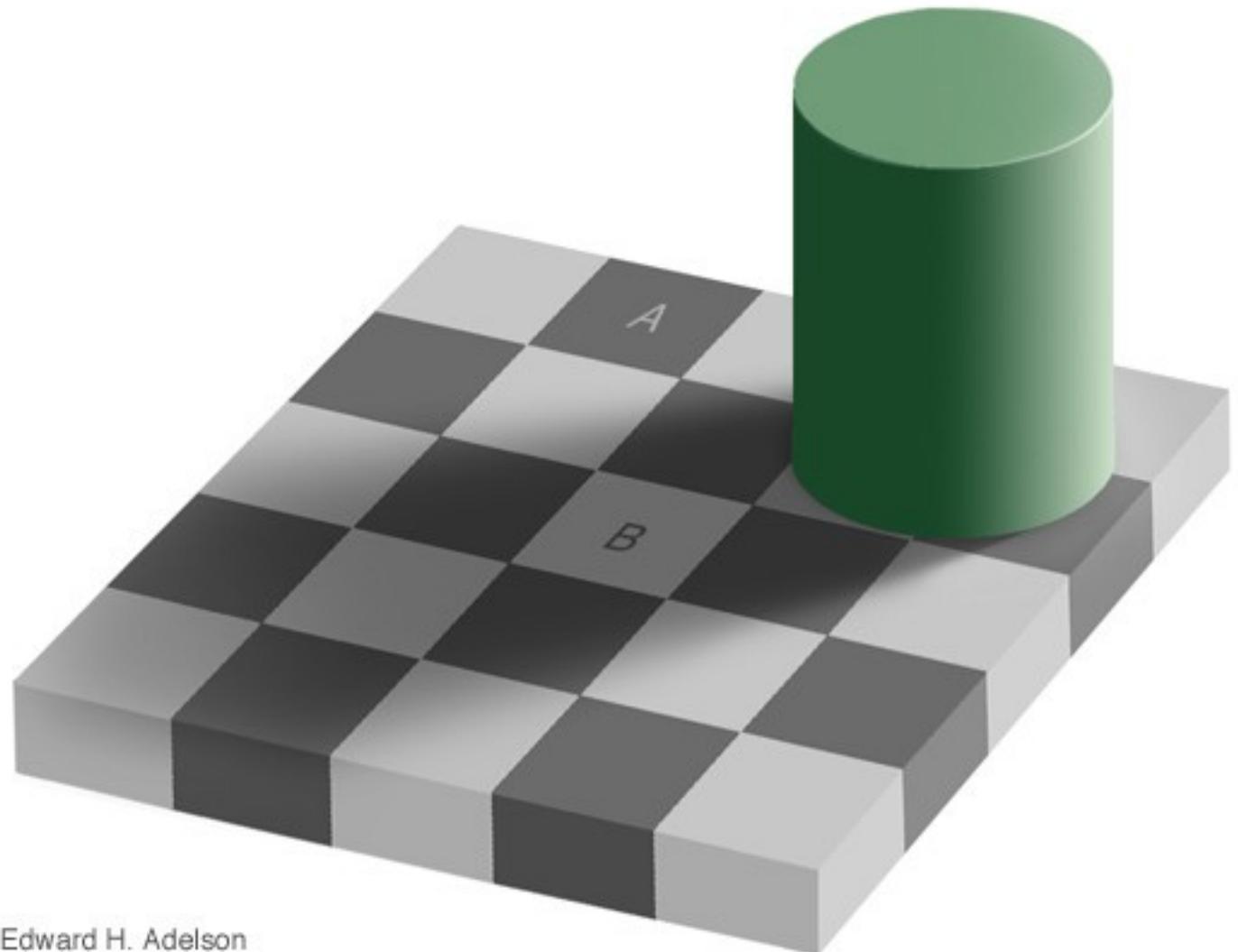
©wikipedia

TI & perception visuelle

Exemple :

l'illusion de l'échiquier.

Plusieurs mécanismes sont en jeu, du très bas niveau (renforcement local des contrastes) au très haut niveau (interprétation de l'ombre et reconnaissance d'un échiquier)



Edward H. Adelson

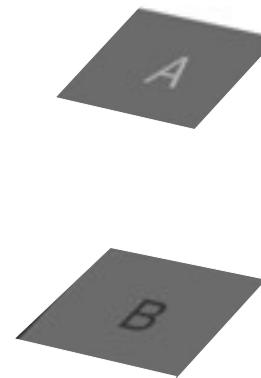
TI & perception visuelle

Checker-shadow illusion:
The squares marked A and B
are the same shade of gray.

Exemple :

l'illusion de l'échiquier.

Plusieurs mécanismes sont en jeu, du très bas niveau (renforcement local des contrastes) au très haut niveau (interprétation de l'ombre et reconnaissance d'un échiquier)

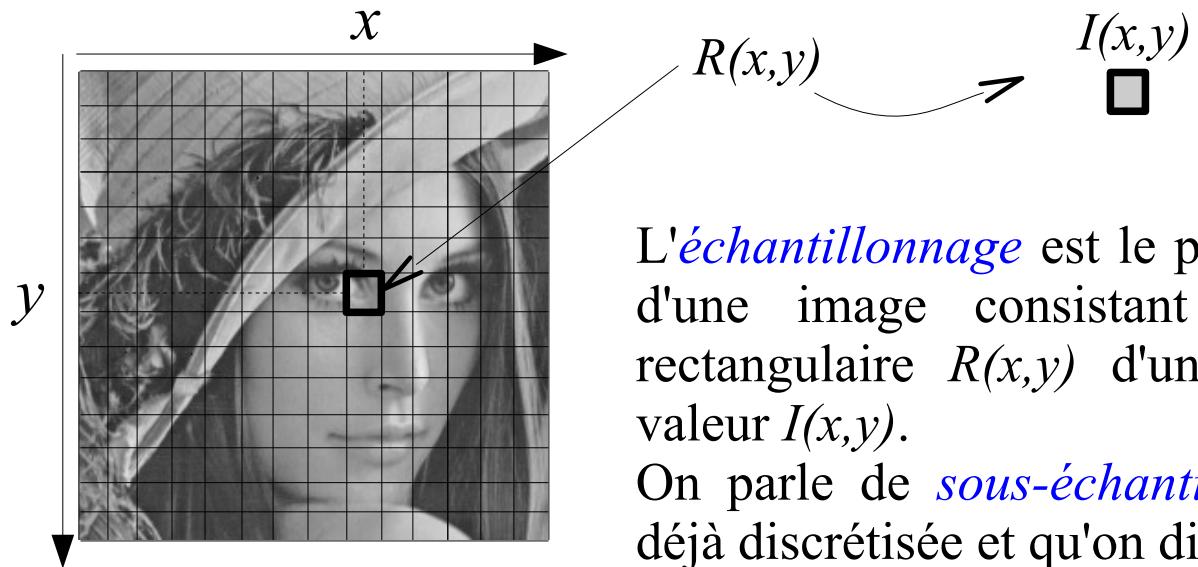


Edward H. Adelson

Types d'images, de capteurs

<i>Phénomène physique</i>	<i>Grandeur mesurée</i>	<i>Capteur</i>
Émission et réflexion de la lumière visible	Réflectivité, luminance,...	CCD, CMOS, Barrettes CCD,...
Rayonnement infra-rouge	Luminance IR (chaleur), ...	Bolomètres,...
Écho ultra sonore	Distance, densité de tissus,...	Échographie, sonar,...
Résonance magnétique	Présence d'un corps chimique,...	IRM, RMN,...
Écho électromagnétique	Distance, specularité de surfaces,..	Radar, SAR,...
Absorption des rayons X	Densité de tissus,...	Radiographie, tomographie,...

Images numériques



L'*échantillonnage* est le procédé de discréétisation spatiale d'une image consistant à associer à chaque zone rectangulaire $R(x,y)$ d'une image continue une unique valeur $I(x,y)$.

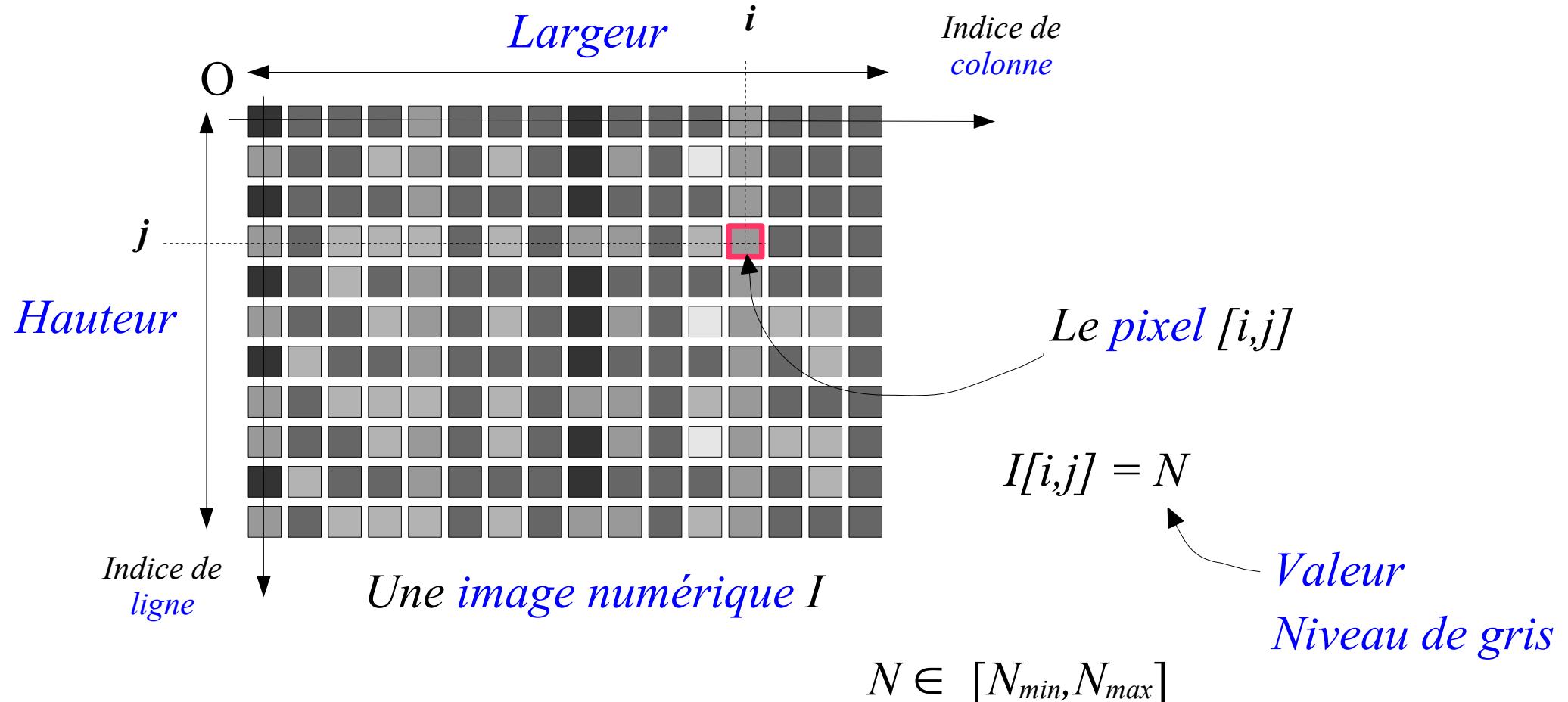
On parle de *sous-échantillonnage* lorsque l'image est déjà discréétisée et qu'on diminue le nombre d'échantillons.

La *quantification* désigne la limitation du nombre de valeurs différentes que peut prendre $I(x,y)$.



Une *image numérique* est une image échantillonnée et quantifiée.

Pixels et niveaux de gris



$$(N_{max} - N_{min}) = \text{nombre de niveaux de gris}$$

$$\log_2(N_{max} - N_{min}) = \text{dynamique}$$

Échantillonnage et quantification

Résolution...

...spatiale :

Échantillonnage



256x256



128x128



64x64



32x32

...tonale :

Quantification



6 bits



4 bits



3 bits



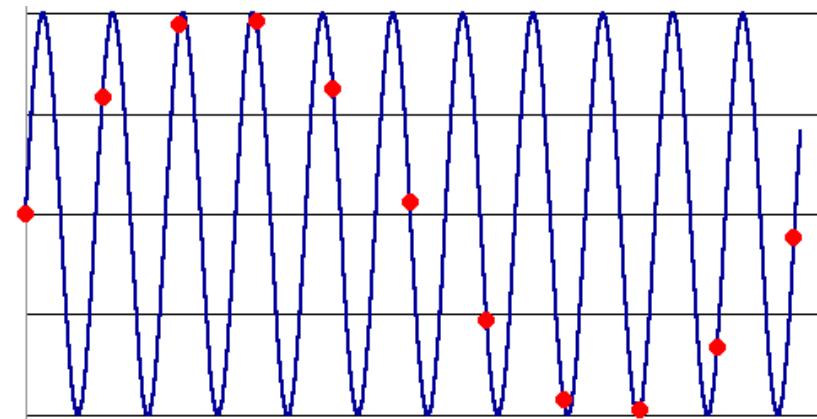
2 bits



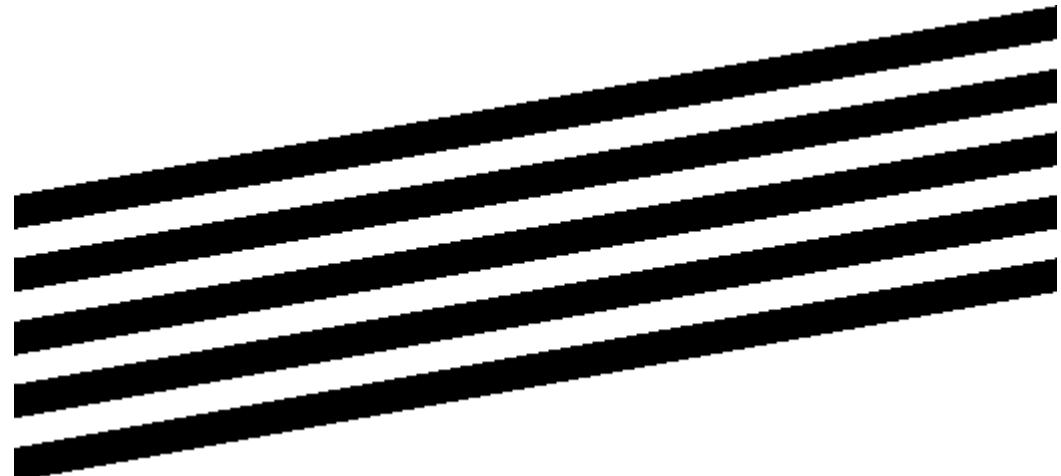
1 bit

Échantillonnage et information

L'échantillonnage est une étape fondamentale qui doit tenir compte du contenu informationnel pertinent de l'image à analyser. Sur l'exemple ci-dessous, en 1d, le signal échantillonné « ressemble » à une sinusoïde de fréquence 12 fois plus faible :

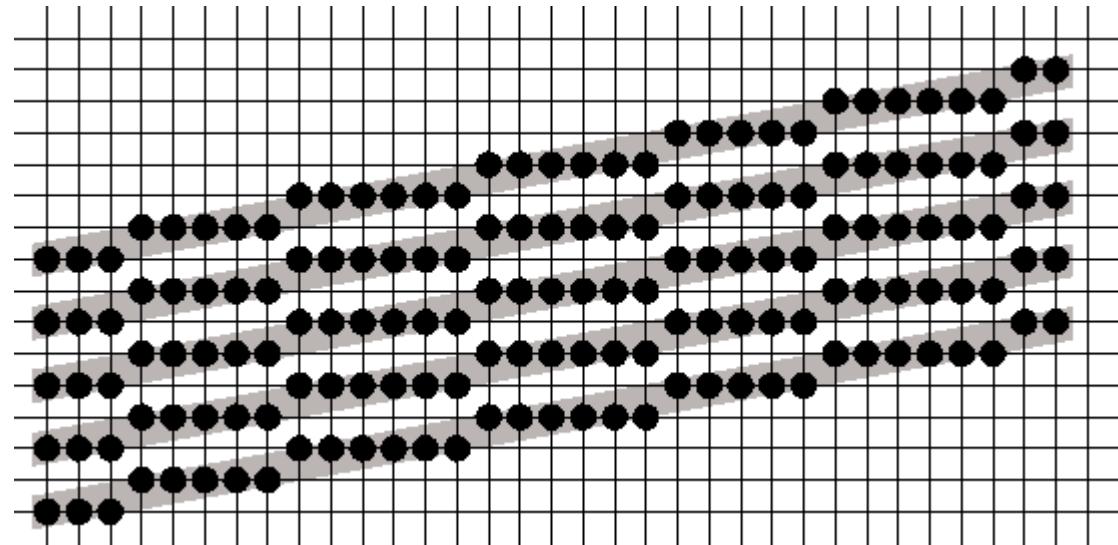


Ce phénomène appelé *aliasing* est encore pire en 2d, car il affecte *la fréquence et la direction* des structures périodiques. Imaginons par exemple qu'on souhaite échantillonner l'image correspondant aux bandes noires ci-contre :

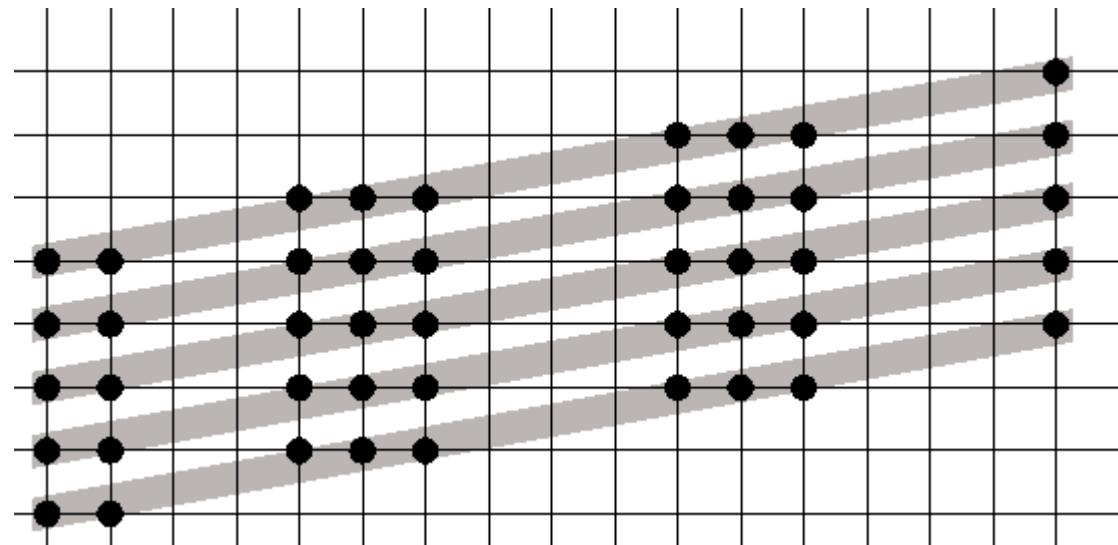


Échantillonnage et information

Avec un échantillonnage adapté, l'image numérique fait apparaître des structures conformes à l'information présente dans l'image :



Mais en considérant seulement 1 échantillon sur 2, une structure différente apparaît, dont l'analyse (ici des bandes verticales, plus épaisses) ne sera pas conforme à la réalité de l'objet :



Échantillonnage et information

Un exemple, sur une image de synthèse :

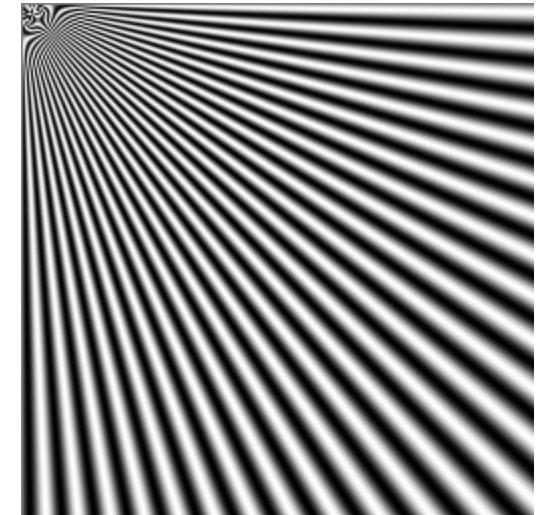


Image originale

Et sur une image naturelle :



Image sous-échantillonnée

Quantification et information

La quantification peut également faire apparaître des distorsions dans les images :

$$I_{quant} = \left\lfloor \frac{I}{n_{quant}} \right\rfloor$$



Comme pour l'échantillonnage, il existe des règles pour déterminer la bonne quantification (le bon nombre de bits) pour coder les images numériques.

L'une dépend du *capteur*, et de sa capacité effective à observer des signaux de valeurs différentes : le *rapport signal sur bruit*.

Le rapport signal sur bruit est défini à partir du rapport entre l'*amplitude des niveaux de gris* mesurables par le capteur ($n_{max} - n_{min}$) et le *niveau du bruit*, en gros l'écart-type s_n de la perturbation aléatoire qui affecte les niveaux de gris. En prenant le logarithme, on a le nombre de bits utile au capteur pour coder les images.

Quantification et information

Outre les capacités du capteur, le nombre de bits réellement nécessaires pour coder une image varie d'une image à l'autre, en fonction de leur *contenu informationnel*.

Ce nombre dépend de l'*entropie*, définie à partir de la distribution des niveaux de gris de l'image (cf plus loin, modèle statistique).



$$E = \sum_{i \leq N} -p_i \log_2(p_i)$$

Où N est le nombre de niveaux de gris présents, p_i est la proportion ($0 < p_i < 1$) de points de l'image ayant pour niveau de gris i . Cette grandeur représente le nombre moyen de bits par pixel nécessaires pour coder toute l'information présente. Elle est utilisée dans les techniques de compression sans perte pour adapter le volume de donnée des images à leur contenu informationnel.

Modèles et outils fondamentaux

Dans la suite, on présente une introduction aux outils d'analyse d'images numériques les plus courants. Pour des raisons didactiques, la présentation est organisée selon les principaux *modèles mathématiques* qui ont été employés pour traiter les images. Ces différents modèles ne sont cependant *ni exclusifs ni cloisonnés*, et la séparation ne sera pratiquement plus visible dans les cours suivants.

Aux différents modèles présentés correspondent un certain nombre d'*outils fondamentaux*, qui se sont révélés au cours du temps plus ou moins incontournables, que ce soit d'un point de vue pratique ou théorique. Citons : la *convolution*, la *transformée de Fourier*, l'*histogramme*, les *pyramides*, la *corrélation*, la *transformée en tout-ou-rien*, les *ondelettes*...

Le modèle linéaire

Dans le modèle linéaire, la structure mathématique sous-jacente est l'*Espace Vectoriel*. Les opérateurs fondamentaux sont ceux qui préservent la structure d'espace vectoriel, c'est-à-dire les *applications linéaires* :

$$\begin{aligned} f(I + J) &= f(I) + f(J) \\ f(\lambda I) &= \lambda f(I) \end{aligned}$$

Pour les images, ces opérateurs correspondent aux *convolutions* :

La convolution :

C'est l'opérateur de base du traitement linéaire des images. Apparue très tôt dans les premiers systèmes d'analyse d'images sous forme empirique et justifiée par des considérations d'implantation, ce n'est que plus tard qu'on a fourni des justifications physiques et fait le lien théorique avec les filtres et le traitement du signal.

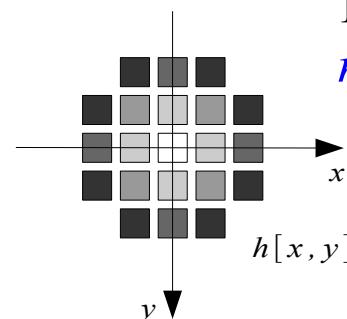
La Convolution

Soit I une image numérique.

Soit h une fonction de $[x_1, x_2] \times [y_1, y_2]$ à valeurs réelles.

La *convolution* de I par h est définie par :

$$(I * h)[x, y] = \sum_{i=x_1}^{x_2} \sum_{j=y_1}^{y_2} h[i, j] \cdot I[x-i, y-j]$$



La fonction h est dite
noyau de convolution

Propriétés de la convolution :

$$\text{COMMUTATIVITÉ } h * g = g * h$$

$$\text{ASSOCIATIVITÉ } (h * g) * k = h * (g * k) = h * g * k$$

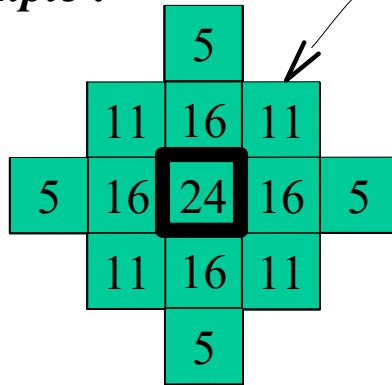
$$\text{DISTRIBUTIVITÉ /+ } h * (g + k) = (h * g) + (h * k)$$

Les nouvelles valeurs du pixel sont calculées par *produit scalaire* entre le noyau de convolution et le *voisinage* correspondant du pixel.

La Convolution

Exemple :

1/152



Pour calculer une convolution, on remplace la valeur de chaque pixel par la valeur du produit scalaire entre les valeurs du noyau de convolution et les valeurs du voisinage du pixel considéré (par rapport à l'origine (0,0) du noyau de convolution).
Attention : implémentation « parallèle ».

Lissage par convolution

Un grand nombre de filtres de lissage peut être obtenu à partir de noyaux de convolution symétriques et normalisés (de somme égale à 1). Voici 3 familles de filtres parmi les plus utilisés :

Moyenne

* *Réponse impulsionnelle :*

$$h(x, y) = \frac{1}{\lambda^2} \text{ si } (x, y) \in [-\lambda/2, +\lambda/2]^2 \\ h(x, y) = 0 \text{ sinon}$$

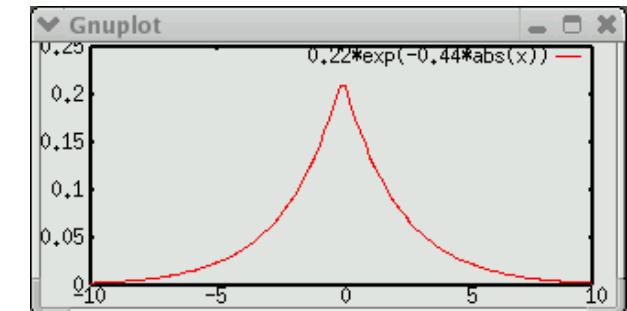
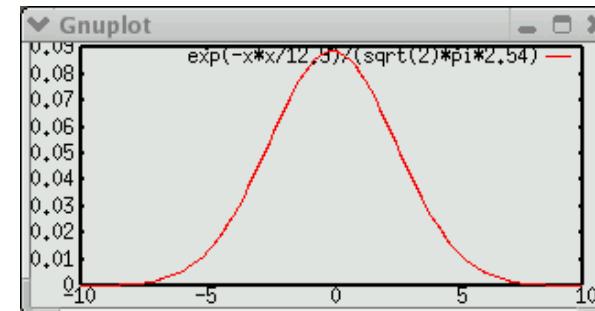
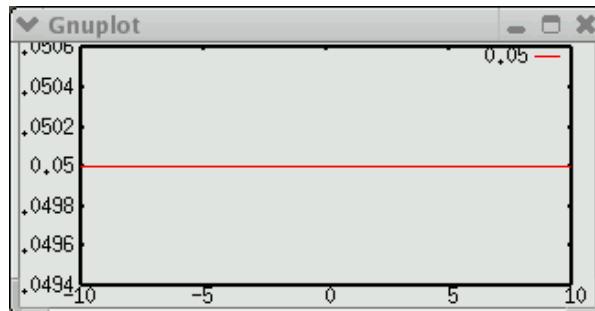
Gauss

$$h(x, y) = \frac{1}{2 \pi \sigma^2} \exp\left(\frac{-(x^2+y^2)}{2 \sigma^2}\right)$$

Exponentiel

$$h(x, y) = \frac{\gamma^2}{4} \exp(-\gamma(|x|+|y|))$$

* *Représentation graphique de la réponse impulsionnelle (en 1d) :*



* *Exemple de noyaux de convolution discrets :*

$$\frac{1}{25} \cdot \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

Filtre moyenneur (5x5)

$$\frac{1}{864} \cdot \begin{pmatrix} 11 & 23 & 29 & 23 & 11 \\ 23 & 48 & 62 & 48 & 23 \\ 29 & 62 & 80 & 62 & 29 \\ 23 & 48 & 62 & 48 & 23 \\ 11 & 23 & 29 & 23 & 11 \end{pmatrix}$$

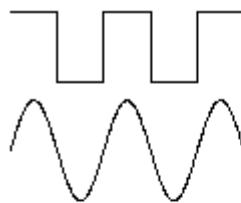
Filtre gaussien ($\sigma = 1,41$)

$$\frac{1}{80} \cdot \begin{pmatrix} 1 & 1 & 3 & 1 & 1 \\ 1 & 3 & 7 & 3 & 1 \\ 3 & 7 & 16 & 7 & 3 \\ 1 & 3 & 7 & 3 & 1 \\ 1 & 1 & 3 & 1 & 1 \end{pmatrix}$$

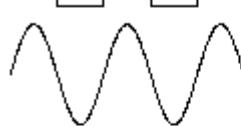
Filtre exponentiel ($\gamma = 0,8$)

Le modèle fréquentiel

Le modèle fréquentiel tend à décrire l'image en termes de *structures périodiques*, en la décomposant *dans une base de fonctions périodiques simples*, comme des sinusoïdes :



- a square wave can be made by adding...



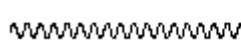
- the fundamental...



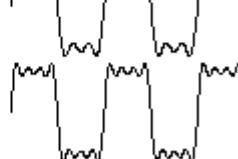
- minus 1/3 of the third harmonic



- plus 1/5 of the fifth harmonic...



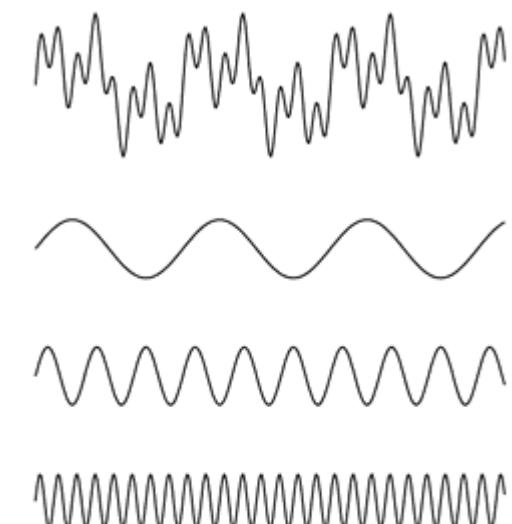
- minus 1/7th of the 7th harmonic...



© **BORES** Signal Processing

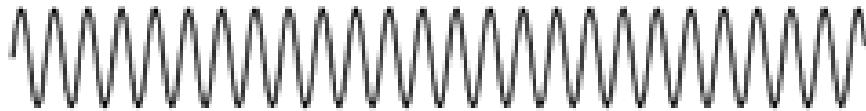
La transformée de Fourier :

Outil fondamental d'analyse en traitement du signal, le pendant bidimensionnel de la TF et sa version discrète peut être appliquée avec profit aux images numériques. Si son utilisation en tant qu'outil analytique et algorithmique s'est estompée en traitement d'images au profit d'approches mieux adaptées à la localisation spatiale des fréquences (ondelettes), elle reste un outil théorique et pédagogique important : la formalisation du phénomène de l'aliasing et des contraintes d'échantillonnage en est un exemple (voir cours Espace d'échelles).



Fourier Analysis: the complex wave at the top can be decomposed into the sum of the three simple waves shown below.

Le modèle fréquentiel



Transformée de Fourier

La transformée de Fourier permet la décomposition d'un signal f en *combinaison linéaire de sinusoïdes complexes*, dont les coefficients $F[u,v]$ dit *coefficients de Fourier*, fournissent des informations sur les *fréquences* (u,v) et permettent des manipulations dans le *domaine fréquentiel*.

Transformée de Fourier discrète bidimensionnelle :

(x,y) sont les coordonnées du *domaine spatial*

(u,v) sont les coordonnées du *domaine fréquentiel*

Directe :

$$F[u, v] = \sum_{x=0}^{w-1} \sum_{y=0}^{h-1} f[x, y] e^{-2i\pi(ux/w + vy/h)}$$

Inverse :

$$f[x, y] = \frac{1}{wh} \sum_{u=0}^{w-1} \sum_{v=0}^{h-1} F[u, v] e^{2i\pi(ux/w + vy/h)}$$

Propriétés de la transformée de Fourier (1) :

ÉCRITURE SOUS FORME MODULE / PHASE

$$F[u, v] = |F[u, v]| e^{i\varphi[u, v]}$$

PÉRIODICITÉ

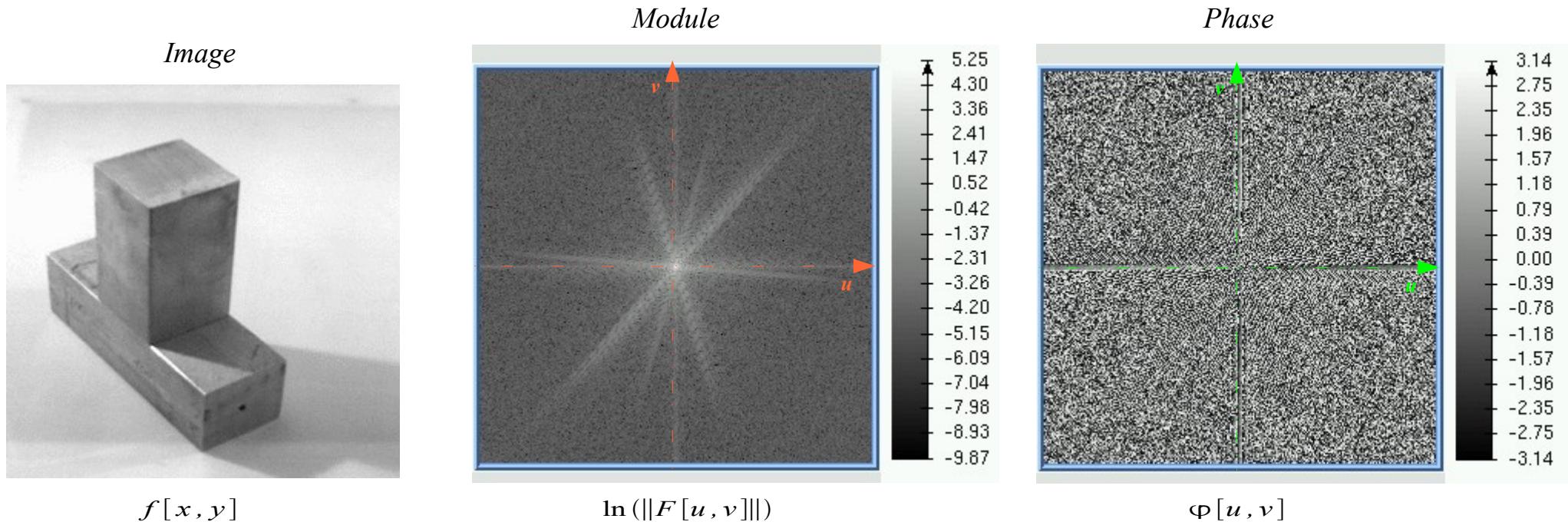
$$F[u, v] = F[u+w, v+h]$$

SYMÉTRIE

Si F est la transformée de Fourier d'une fonction réelle f :

$$F[u, v] = \overline{F[-u, -v]} \quad \text{et donc : } |F[u, v]| = |F[-u, -v]| \quad \text{et} \quad \varphi[u, v] = -\varphi[-u, -v]$$

Transformée de Fourier



Propriétés de la transformée de Fourier (2) :

$$si \begin{cases} f[x,y] & \xrightarrow{\text{TF}} F[u,v] \\ f_1[x,y] & \xrightarrow{} F_1[u,v] \\ f_2[x,y] & \xrightarrow{} F_2[u,v] \end{cases}$$

CORRESPONDANCE CONVOLUTION / PRODUIT

$$\begin{aligned} f_1[x,y]*f_2[x,y] &\rightarrow F_1[u,v] \cdot F_2[u,v] \\ f_1[x,y] \cdot f_2[x,y] &\rightarrow F_1[u,v]*F_2[u,v] \end{aligned}$$

DÉRIVATION

$$\frac{\partial f[x,y]}{\partial x} \rightarrow iuF[u,v] \text{ et } \frac{\partial f[x,y]}{\partial y} \rightarrow ivF[u,v]$$

LINEARITÉ

$$a \cdot f_1[x,y] + b \cdot f_2[x,y] \rightarrow a \cdot F_1[u,v] + b \cdot F_2[u,v]$$

TRANSLATIONS SPATIALES / FRÉQUENTIELLES

$$\begin{aligned} f[x-x',y-y'] &\rightarrow F[u,v] \cdot e^{-2i\pi(ux'/w+vy'/h)} \\ f[x,y] \cdot e^{2i\pi(u'x/w+v'y/h)} &\rightarrow F[u-u',v-v'] \end{aligned}$$

THÉORÈME DE PARSEVAL

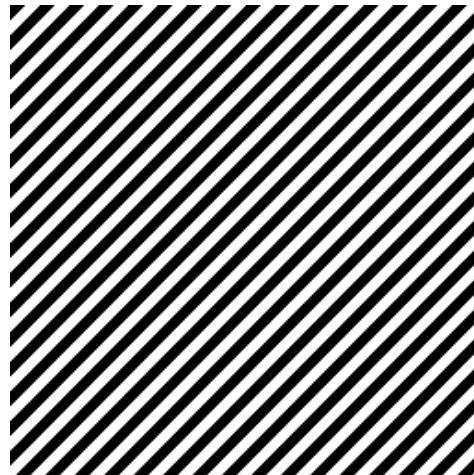
$$\sum_{x=0}^{w-1} \sum_{y=0}^{h-1} \|f[x,y]\|^2 = \frac{1}{wh} \sum_{u=0}^{w-1} \sum_{v=0}^{h-1} \|F[u,v]\|^2$$

QUIZZ - Transformée de Fourier

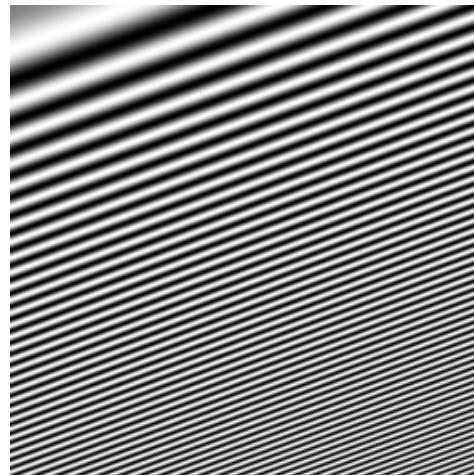
Attribuez à chaque image son spectre de Fourier



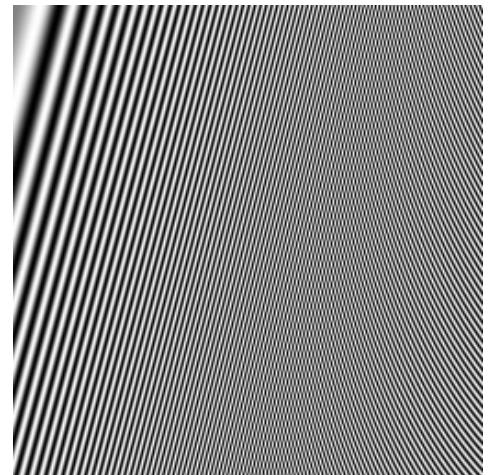
(1)



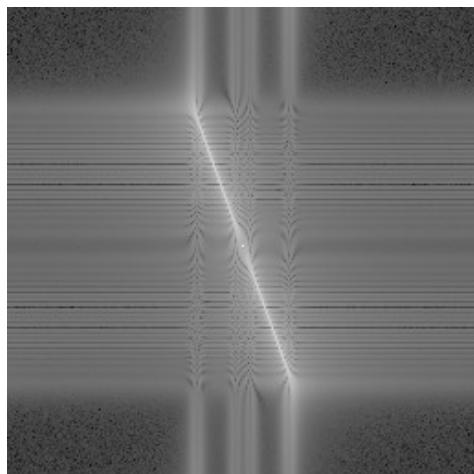
(2)



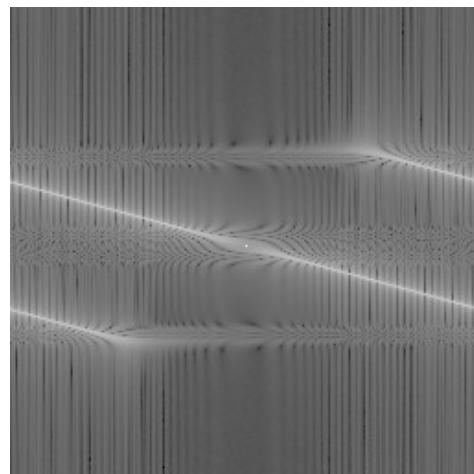
(3)



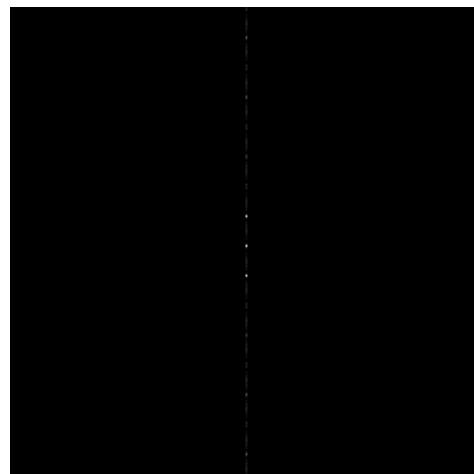
(4)



(a)



(b)



(c)



(d)

QUIZZ - Transformée de Fourier

Attribuez à chaque image son spectre de Fourier



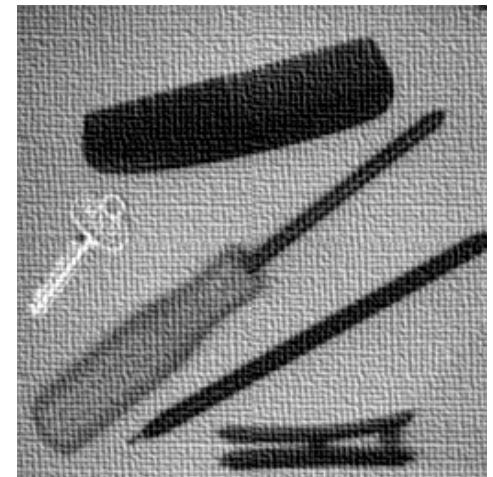
(1)



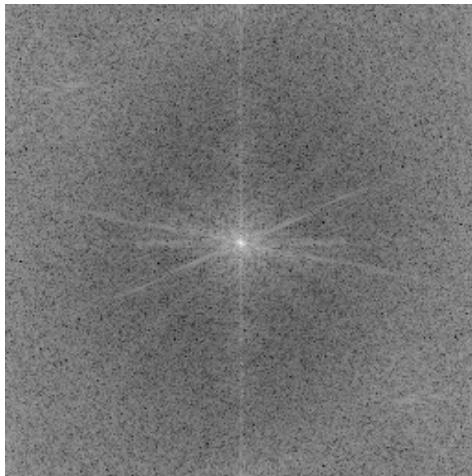
(2)



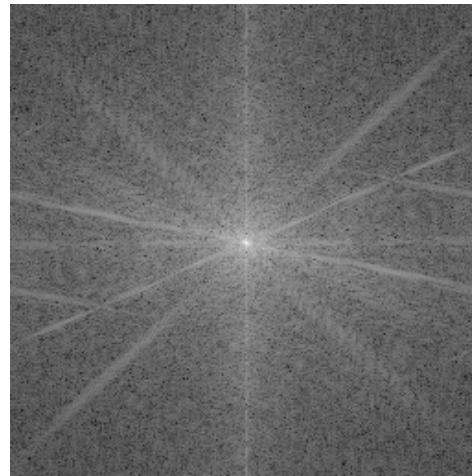
(3)



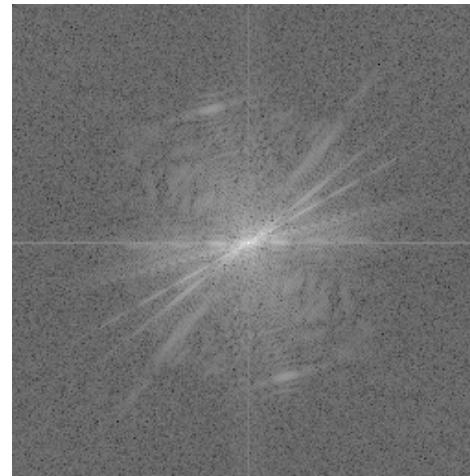
(4)



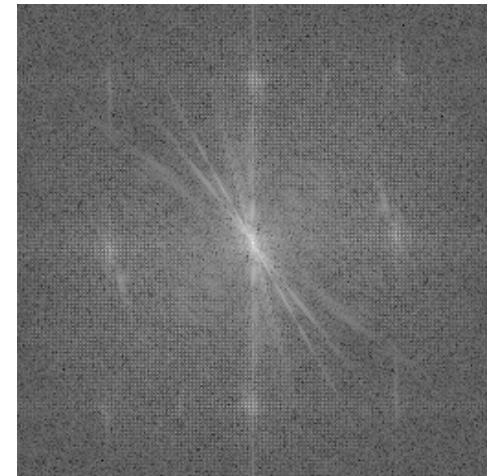
(a)



(b)



(c)



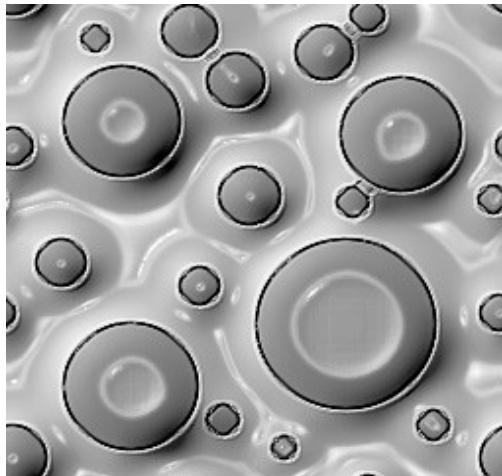
(d)

QUIZZ - Transformée de Fourier

Attribuez à chaque image son spectre de Fourier (CC 2014)



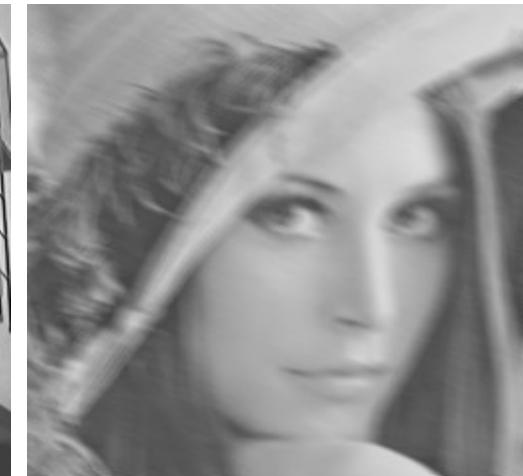
(1)



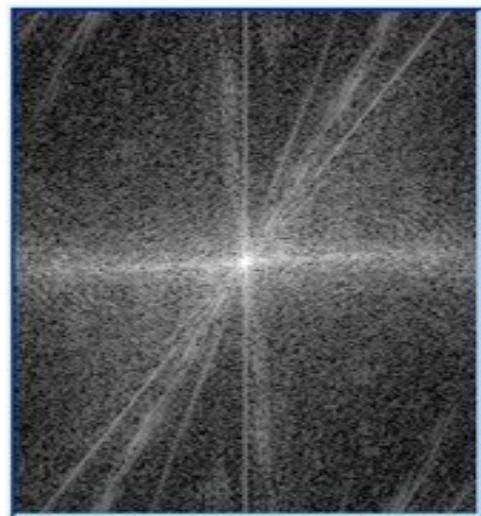
(2)



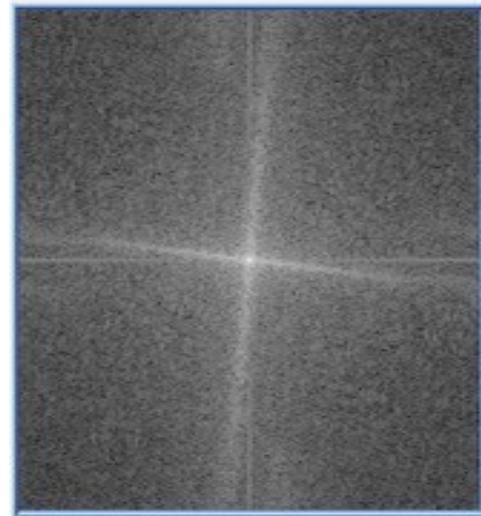
(3)



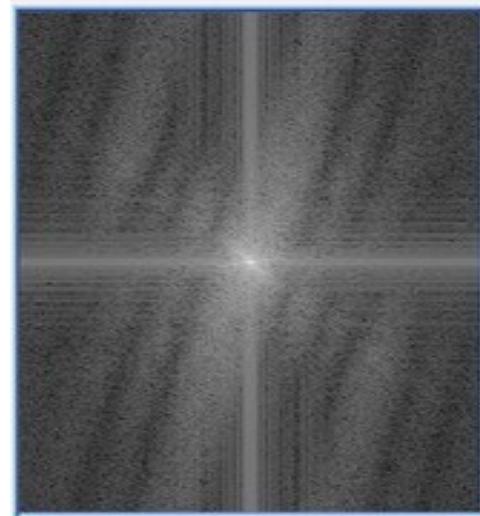
(4)



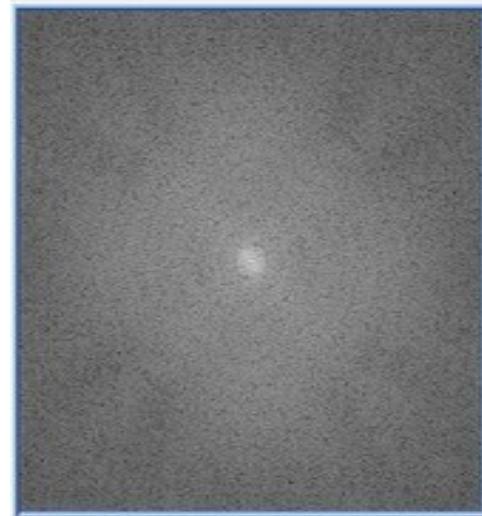
(a)



(b)



(c)

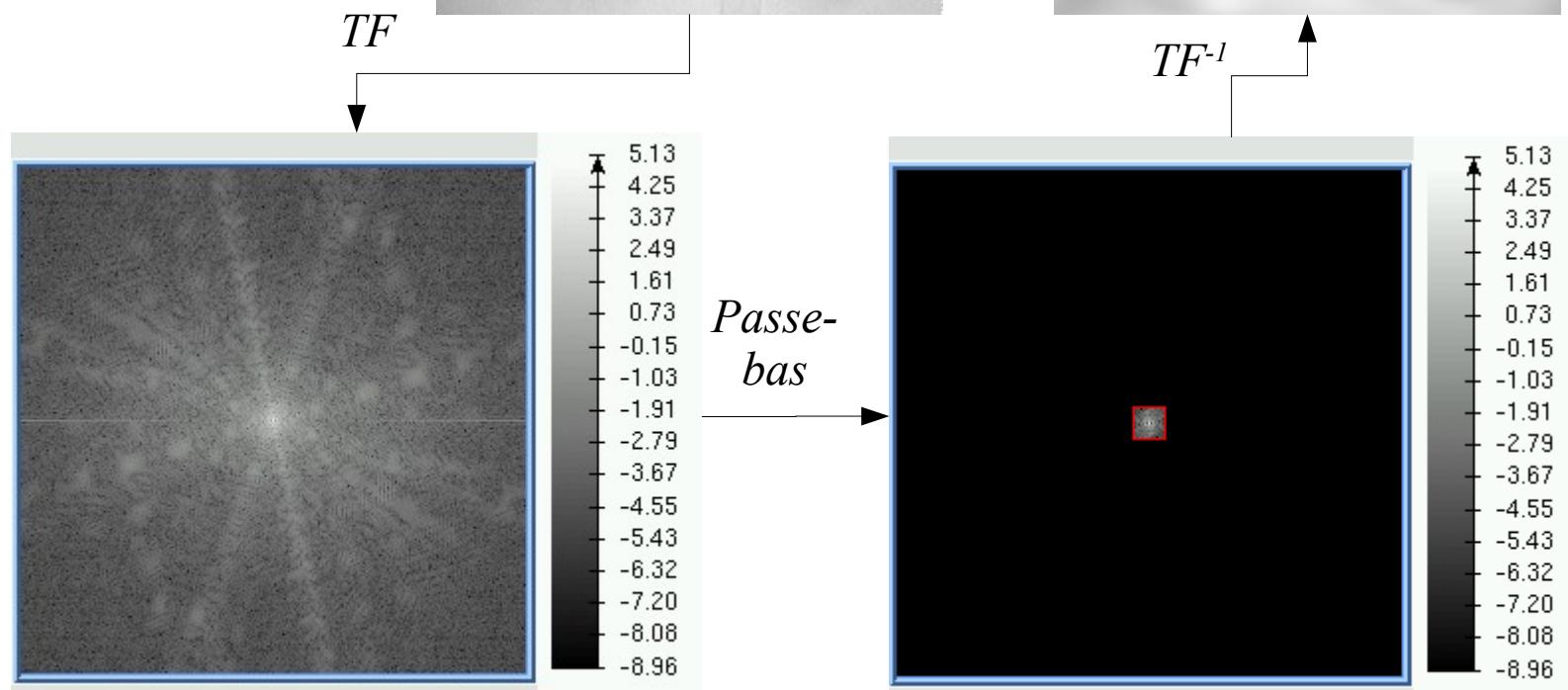


(d)

Filtrage dans le domaine de Fourier (1)

Filtrage passe-bas

Le filtrage passe-bas est la multiplication dans le domaine fréquentiel par une *fonction porte* (fonction indicatrice d'un intervalle $[-u_{max}, u_{max}] \times [-v_{max}, v_{max}]$).



Filtrage dans le domaine de Fourier (2)

Filtrage coupe-bande

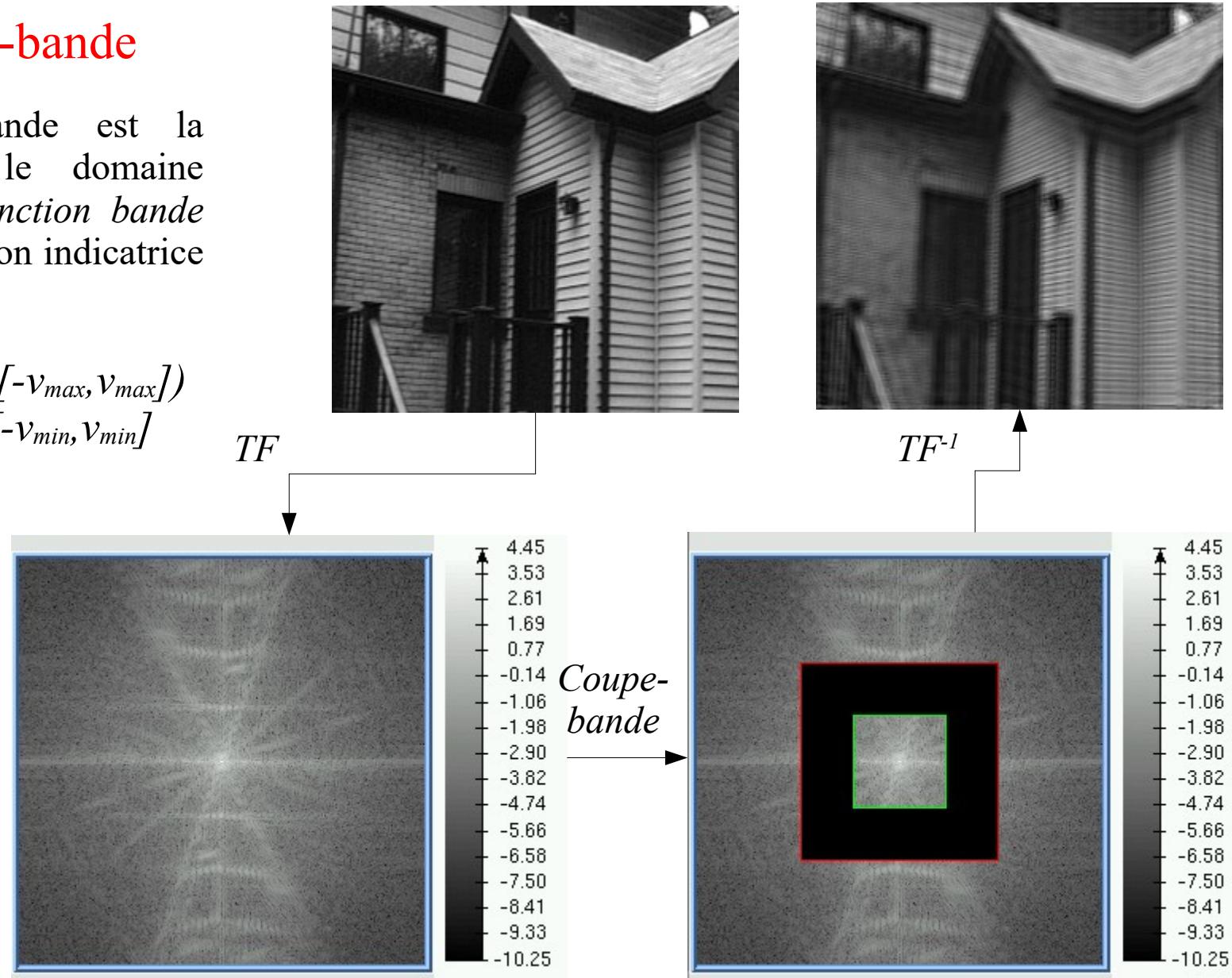
Le filtrage coupe-bande est la multiplication dans le domaine fréquentiel par une *fonction bande complémentaire*, fonction indicatrice de l'ensemble :

$$\begin{aligned} & (\mathbb{R}^2 \setminus [-u_{max}, u_{max}] \times [-v_{max}, v_{max}]) \\ & \cup [-u_{min}, u_{min}] \times [-v_{min}, v_{min}] \end{aligned}$$

Notons que dans ce cas comme le précédent, la valeur de la fréquence origine $F[0,0]$ est inchangée. Or :

$$F[0,0] = \sum_{x=0}^w \sum_{y=0}^h f[x, y]$$

La somme des niveaux de gris dans le domaine spatial reste donc constante.

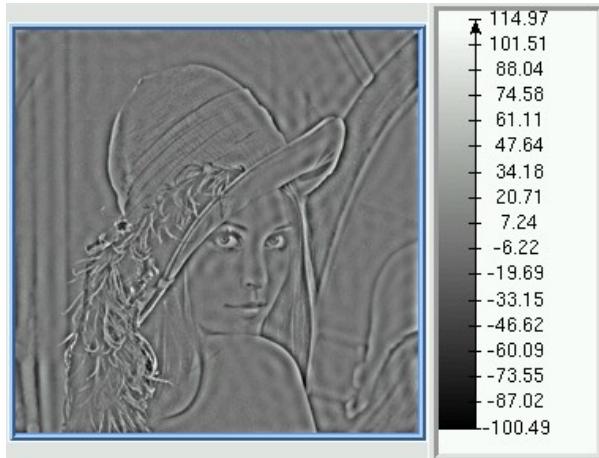


Filtrage dans le domaine de Fourier (3)

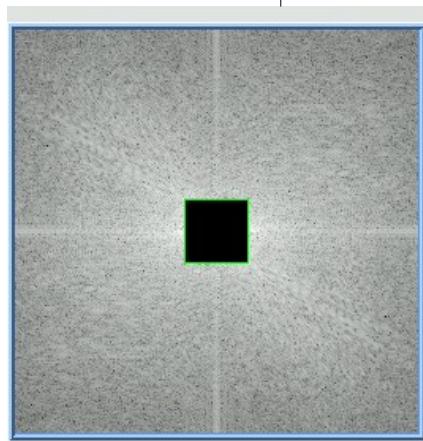
Le filtrage passe-haut correspond à la multiplication dans le domaine fréquentiel d'une fonction porte-complémentaire.

Filtrage passe-haut

Image originale

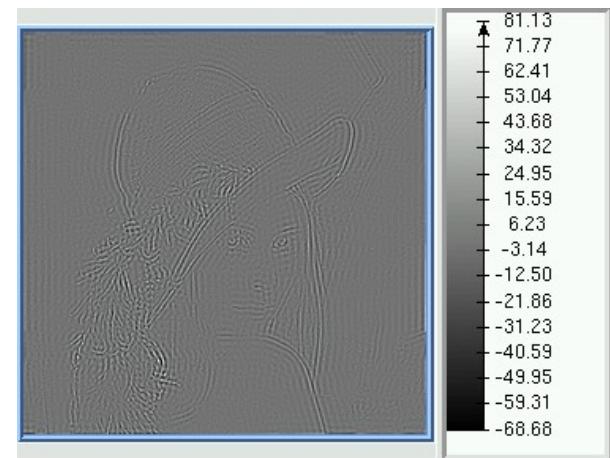
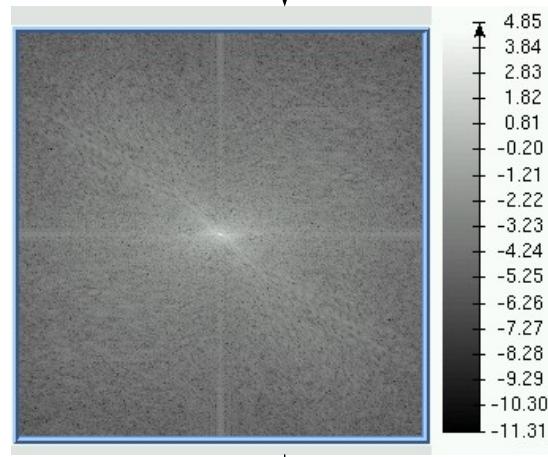


$\uparrow \text{TF}^{-1}$

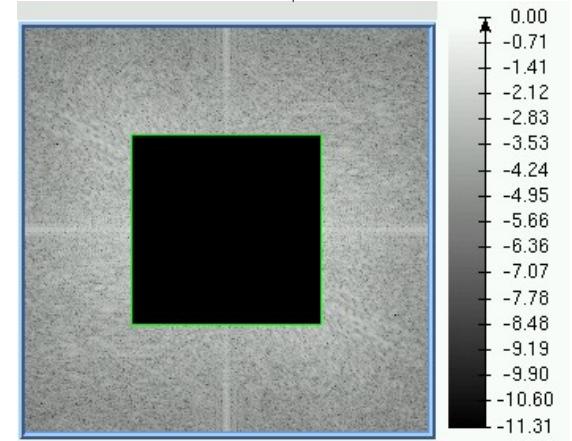


Filtre passe-haut $\phi = 20$

$\downarrow \text{TF}$



$\uparrow \text{TF}^{-1}$



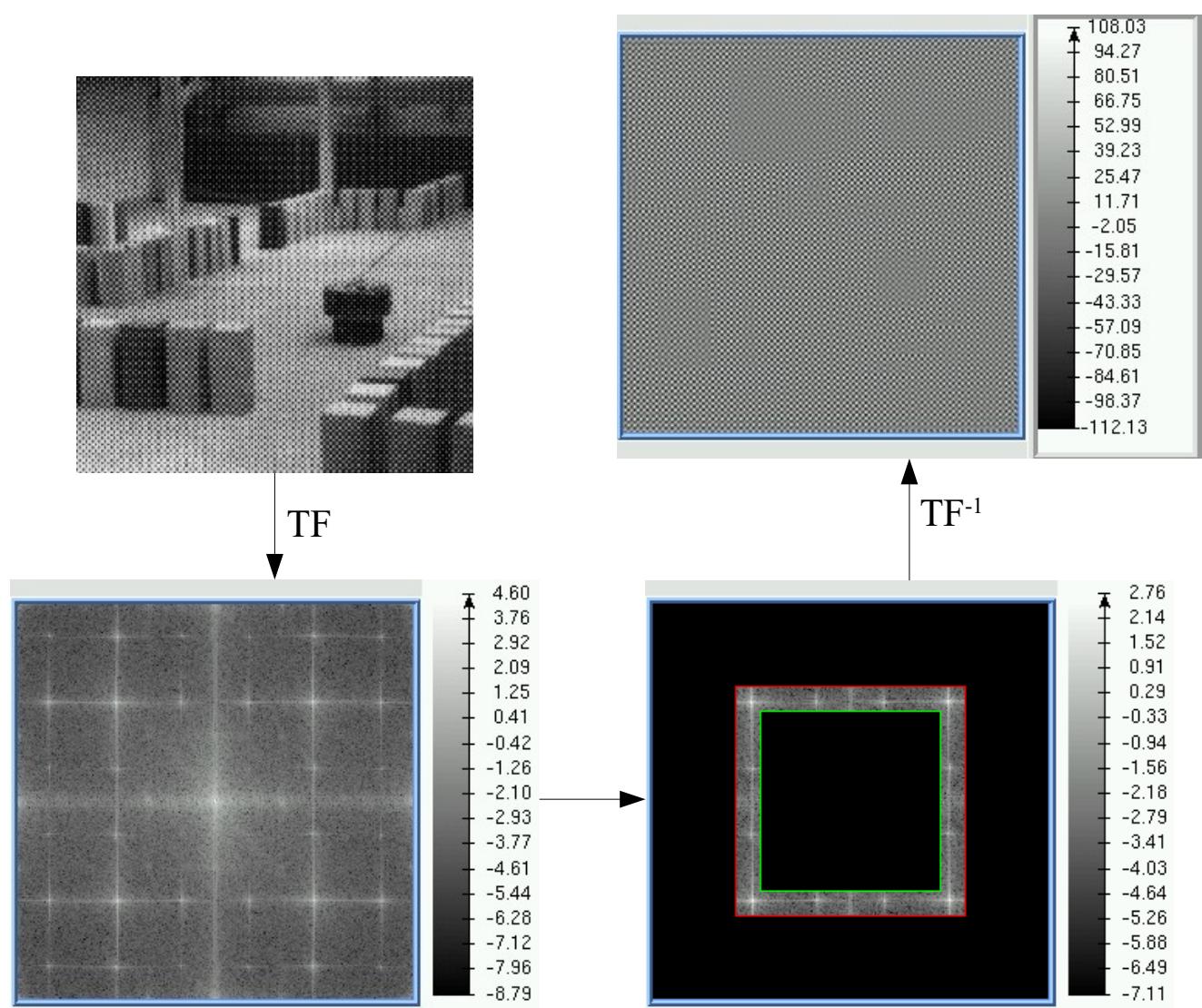
Filtre passe-haut $\phi = 60$

Filtrage dans le domaine de Fourier (4)

Le filtrage passe-bande correspond à la multiplication dans le domaine fréquentiel par une fonction bande symétrique.

Dans ce cas comme dans le précédent, la valeur de la fréquence origine $(0,0)$ est annulée. Par conséquent, l'image correspondante dans le domaine spatial est de somme nulle, elle comporte donc des valeurs négatives.

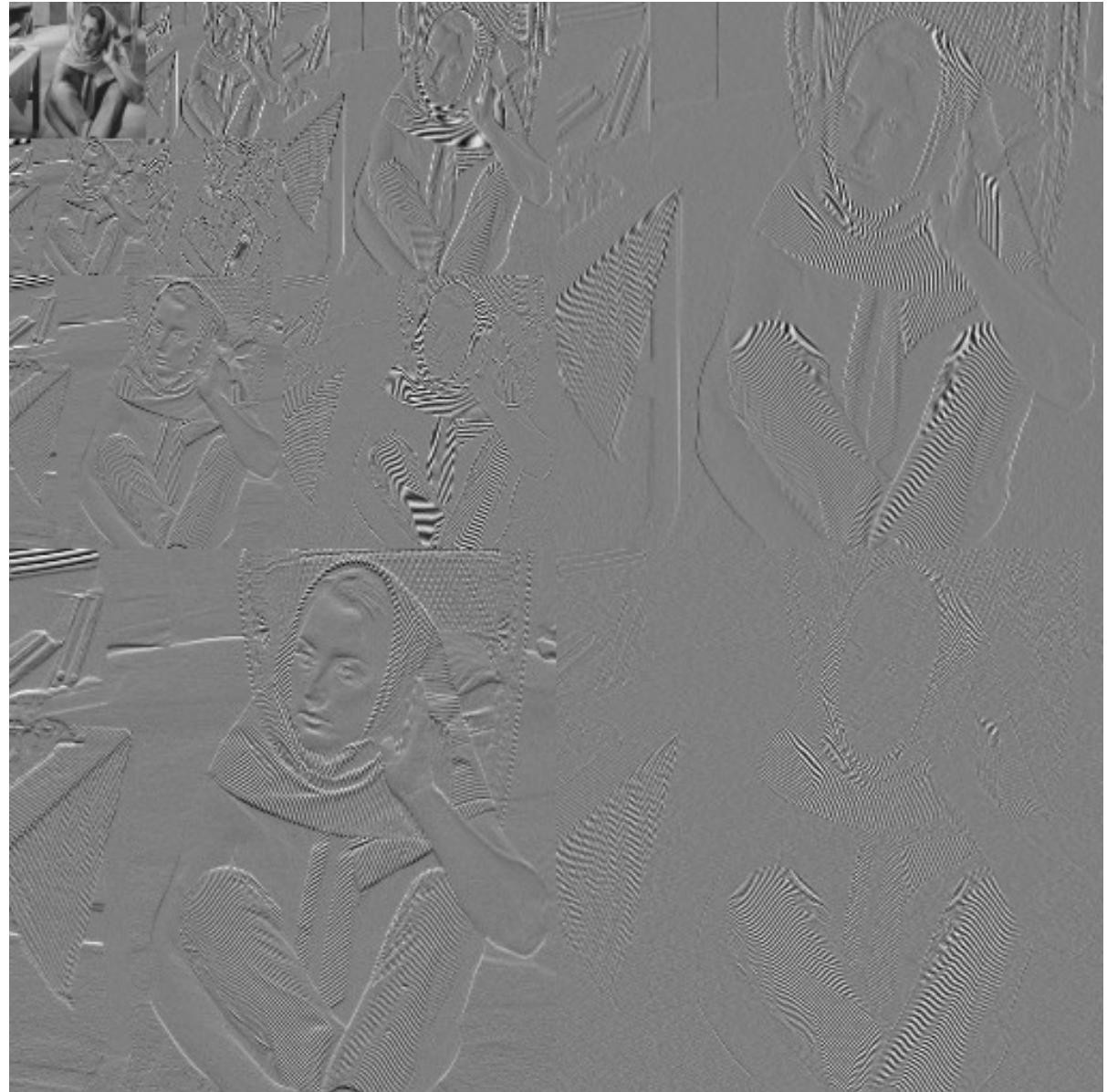
Filtrage passe-bande



Transformées en ondelettes

Pour pallier les limitations de la transformée de Fourier en analyse d'images, on utilise plus souvent des représentations **espace × fréquence**, telles que les transformées en ondelettes, qui fournissent une **décomposition en fréquence spatialement localisée**.

Ex ci-contre : décomposition en ondelettes de niveau 3 sur l'image « Barbara »



Transformées en ondelettes

Les transformées en ondelettes sur les images se calculent généralement de façon efficace, par l'utilisation d'un couple de filtre 1d (L, H) complémentaires, où $I \star L$ est une approximation de l'image (passe-bas) et $I \star H$ la partie résiduelle (passe haut).

Le couple $(I \star L, I \star H)$ doit de plus permettre de **reconstruire** l'image originale, et constitue donc un **codage**, au même titre que la transformée de Fourier.



Transformées en ondelettes

Par exemple, l'ondelette de Haar est définie par le couple de convolutions 1d (**analyse**) :

$$(I * L)[x] = (I[2x] + I[2x+1])/2$$
$$(I * H)[x] = (I[2x+1] - I[2x])/2$$

La reconstruction (**synthèse**) est simplement donnée par :

$$I[2x] = (I * L)[x] - (I * H)[x]$$
$$I[2x+1] = (I * L)[x] + (I * H)[x]$$

Ci-contre : Application du couple de filtre (L,H) sur les lignes de l'image « Barbara »



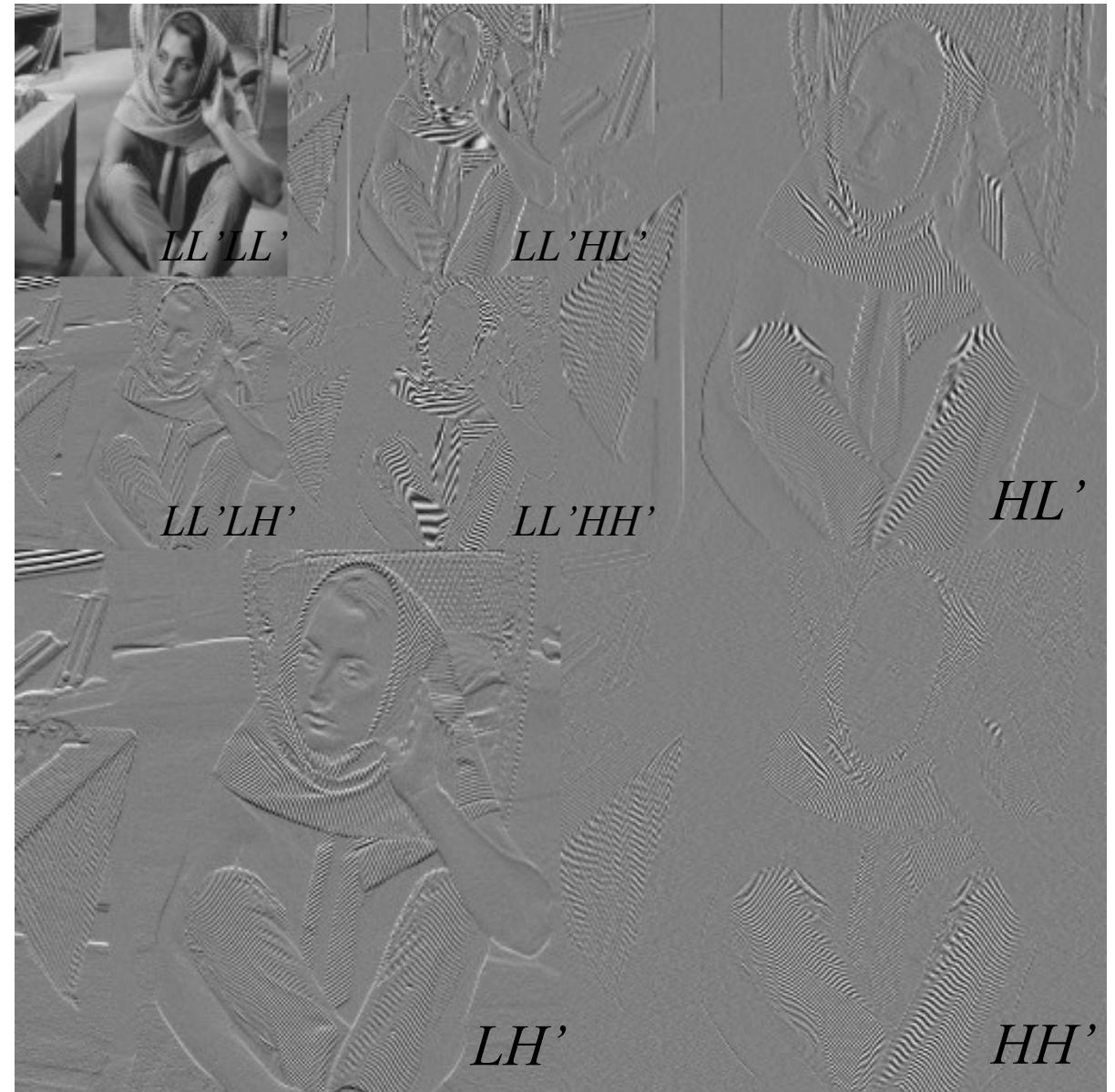
Transformées en ondelettes

Ensuite le même couple de filtres transposés (L', H') est appliqué sur les colonnes :



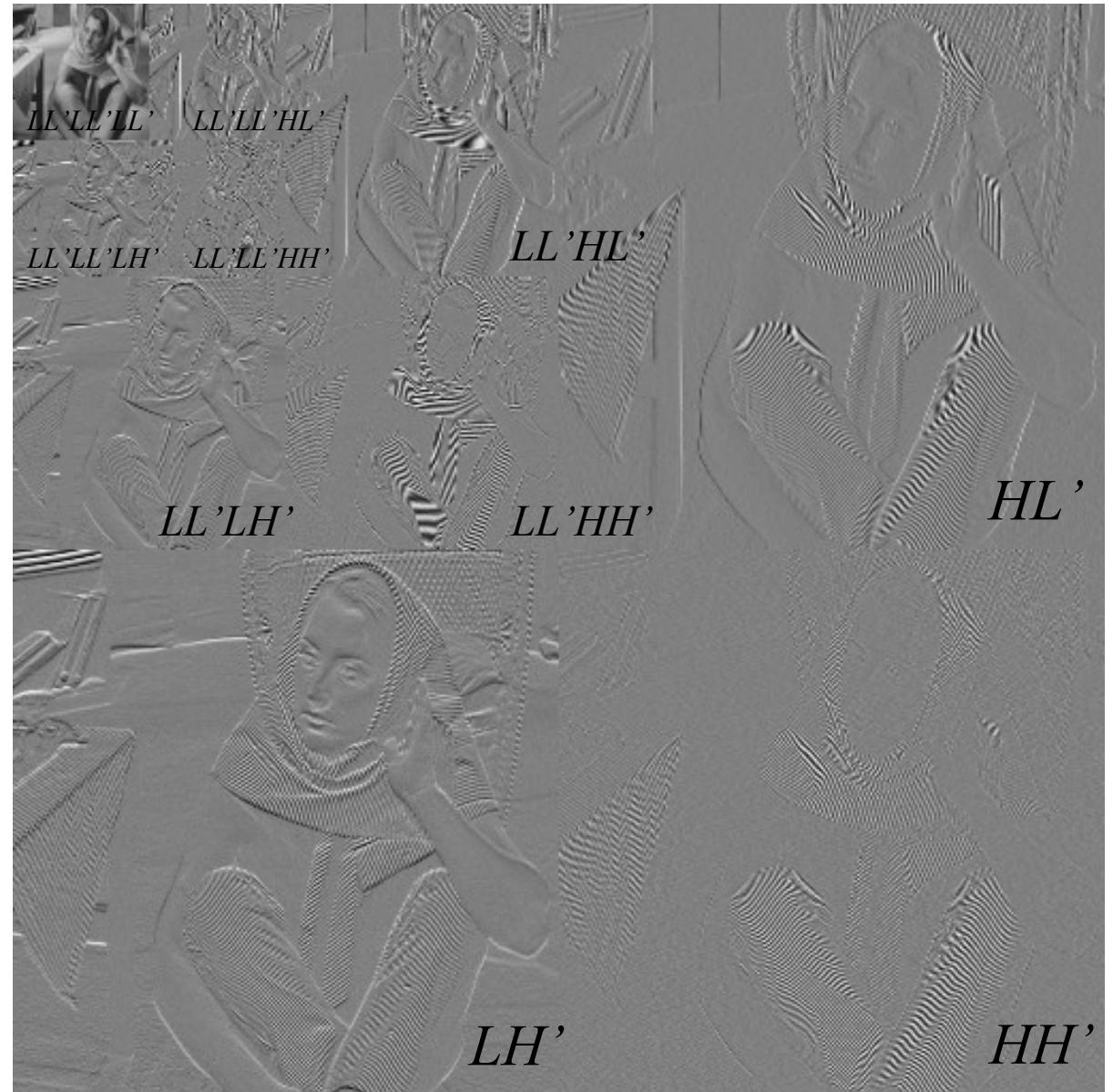
Transformées en ondelettes

Le même procédé est appliqué récursivement sur l'image d'approximation LL' (niveau 2) :



Transformées en ondelettes

Et ainsi de suite (niveau 3)...

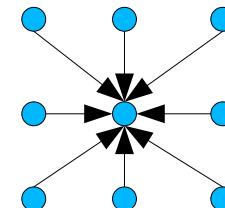
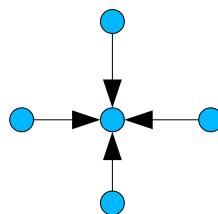


Le modèle statistique

On s'intéresse dans ce modèle aux propriétés statistiques des images : la distribution des valeurs prises par les pixels, la corrélation existant entre des pixels spatialement proches, la fréquence d'occurrence de certaines structures spatiales...

Les mesures statistiques fournissent des grandeurs et fonctions empiriques sur lesquelles peuvent s'appuyer des modèles probabilistes utilisés par les algorithmes d'analyse d'images.

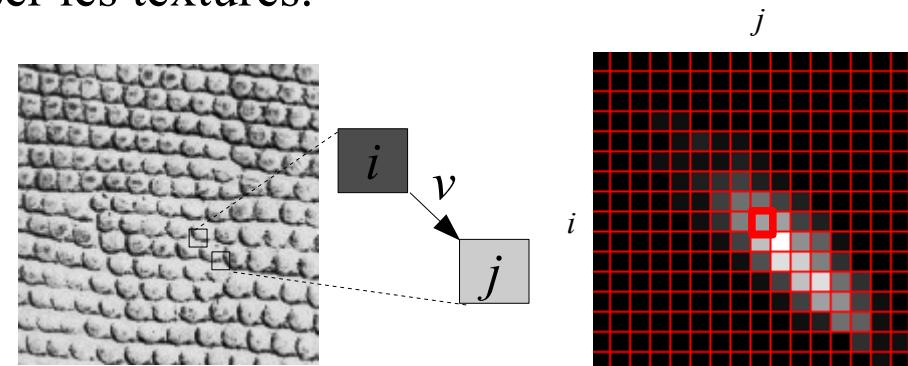
Par exemple, le modèle des champs de Markov considère l'image comme la réalisation d'un champ aléatoire (chaque pixel correspondant à une variable aléatoire), où la valeur prise par un pixel ne dépend que de celle de ses voisins (selon une topologie discrète donnée, voir plus loin).



Le modèle statistique

Un autre exemple remarquable d'analyse statistique est l'utilisation des matrices de co-occurrence, habituellement utilisées pour caractériser les textures.

La matrice de cooccurrence M_v associée au vecteur v , est la matrice de taille $N \times N$ (N est le nombre de niveaux de gris), tel que $M_v(i,j)$ représente la fréquence du couple de valeurs (i,j) parmi les couples de pixels $(x,x+v)$.

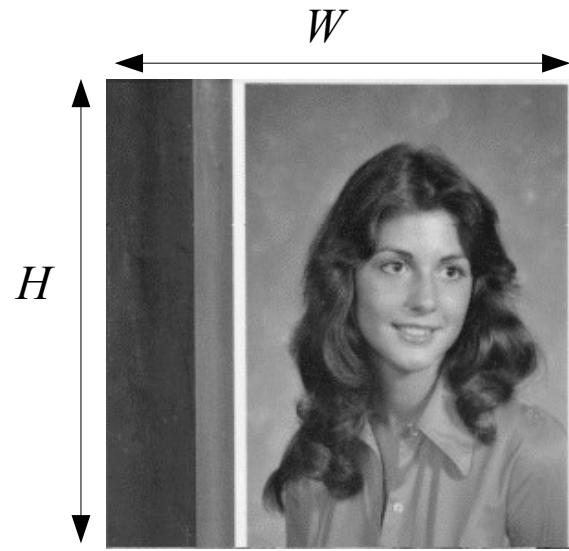


Ces techniques, et d'autres exemples de modèles statistiques seront présentées plus en détail dans les prochains cours. Dans le cadre de cette introduction, nous développerons seulement le premier outil statistique d'analyse des images : l'histogramme.

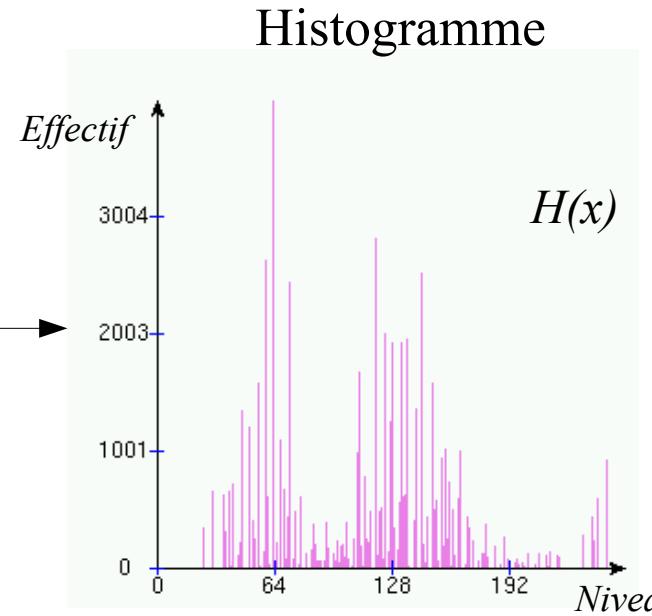
L'histogramme :

Outil de base pour l'étude des capteurs ou de la dynamique d'une scène, il est utilisé par certains opérateurs d'analyse. On retiendra cependant qu'il ne faut pas considérer l'histogramme comme une caractéristique fondamentale de l'image dans la mesure où on peut le transformer radicalement sans changer significativement l'image.

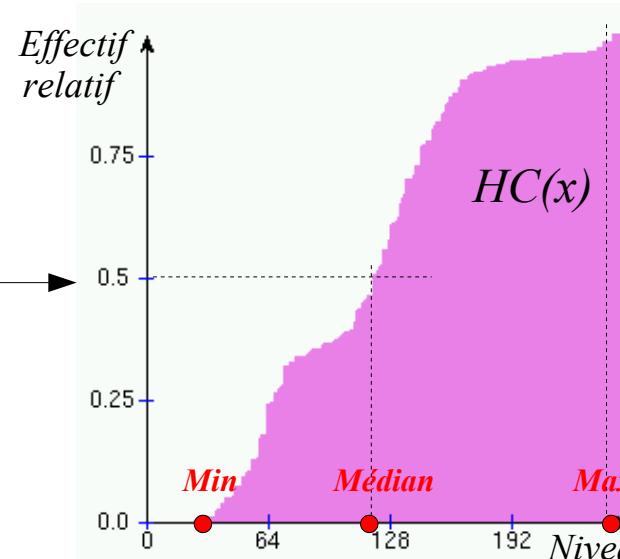
Histogramme



L'histogramme représente la répartition des pixels en fonction de leur niveau de gris. Il fournit diverses informations comme les statistiques d'ordre (voir ci-contre), l'entropie (voir précédemment), et peut permettre d'isoler des objets.



Histogramme cumulatif normalisé



Traitement à base d'histogramme

On présente dans la suite quelques traitement d'analyse effectués uniquement à partir de l'histogramme. Retenons que certains de ces traitements sont souvent calculés au niveau des capteurs, et qu'en général leur pertinence est très intimement liée aux conditions d'acquisition.

(1) Normalisation

→ exploiter toute la dynamique de codage.

(2) Égalisation

→ équilibrer la dynamique de codage et augmenter le contraste.

(3) Segmentation

→ simplifier l'image en regroupant les pixels selon leurs valeurs.

Histogramme : normalisation

La *normalisation d'histogramme*, ou *expansion de dynamique*, est une *transformation affine* du niveau de gris des pixels de telle sorte que l'image utilise *toute la dynamique* de représentation.

D : dynamique
Nmin : la plus petite valeur dans l'image
Nmax : la plus grande valeur dans l'image

$$f_{\text{new}}[x, y] = (f[x, y] - N_{\min}) \cdot \frac{2^D - 1}{N_{\max} - N_{\min}}$$

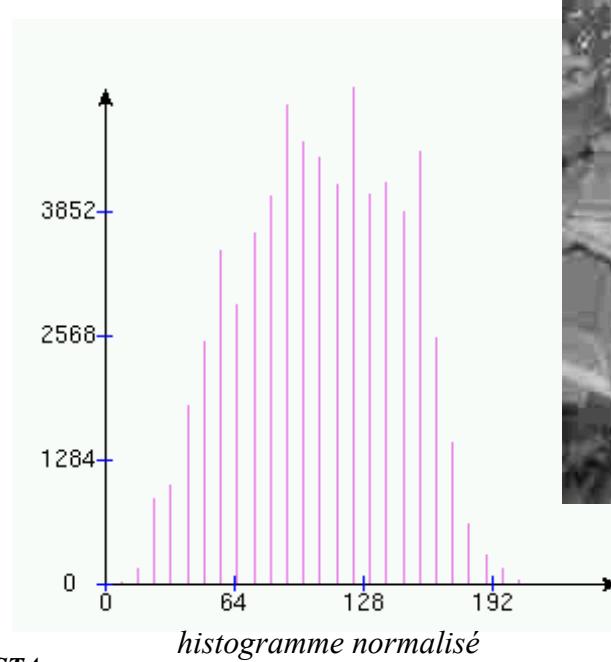
Pour rendre la normalisation moins sensible aux valeurs marginales (outliers), on utilise généralement un paramètre β , $0 < \beta < 1$, et on prend :

$$N_{\min} \in HC^{-1}(\beta)$$

$$N_{\max} \in HC^{-1}(1 - \beta)$$



histogramme d'origine



histogramme normalisé



expansion de dynamique
 $f_{\text{new}}[x, y]$

Histogramme : égalisation

L'*égalisation d'histogramme* est une transformation des niveaux de gris dont le principe est d'*équilibrer* le mieux possible la distribution des pixels dans la dynamique (Idéalement, on cherche à obtenir un *histogramme plat*).

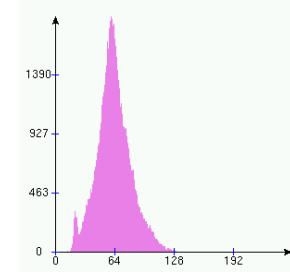
La technique classique consiste à rendre « le plus linéaire possible » l'histogramme cumulé de l'image en utilisant la transformation suivante :

$$f_{new}[x, y] = (2^D - 1) \cdot \frac{HC(f[x, y])}{wh}$$

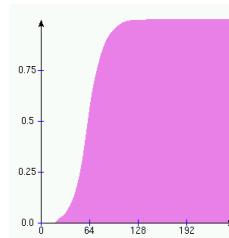
D : dynamique
 (w, h) : dimension de l'image
 $HC(\cdot)$: histogramme cumulé



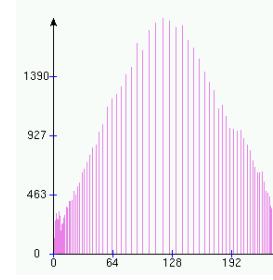
Original $f[x, y]$



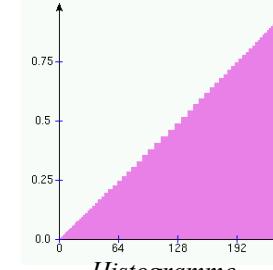
Histogramme de f



Histogramme cumulé de f



Histogramme de f_{new}



Histogramme cumulé de f_{new}



Après égalisation $f_{new}[x, y]$

Le résultat est une *augmentation globale du contraste* dans l'image. Notez dans l'exemple ci-dessus l'accentuation des défauts avec la mise en évidence du bruit spatial fixe (effet de tramage) de l'imageur infra-rouge.

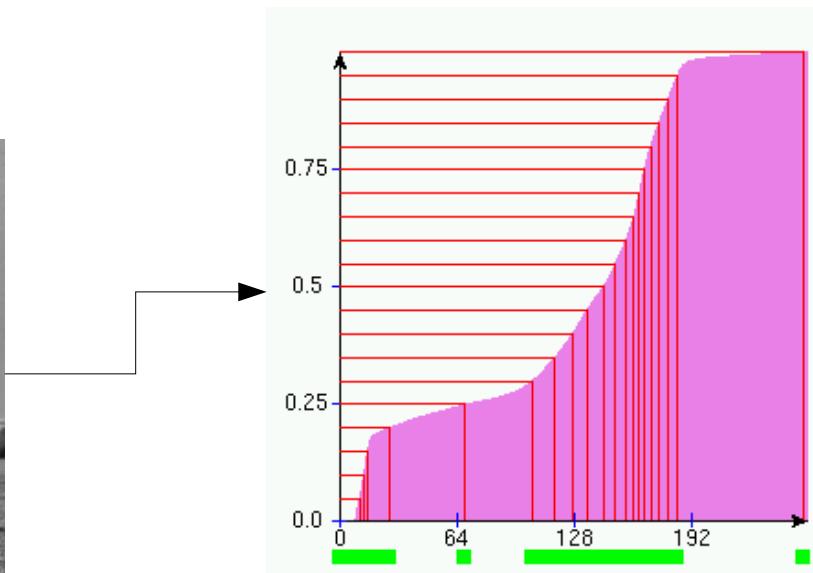
Histogramme : segmentation

Il existe des techniques de segmentation basées sur un regroupement des niveaux de gris à partir de l'histogramme. Ces techniques sont rarement efficaces car elles ne considèrent que la valeur des pixels sans tenir compte de critères géométriques ou topologiques (voir cours Segmentation).

Par exemple, la méthode ci-dessous calcule un certain nombre de quantiles à partir de l'histogramme cumulé, les regroupe par classes en fonction d'un critère de distance, puis attribut la même étiquette aux pixels dont la valeur est la plus proche d'une classe donnée :



Image originale



Histogramme cumulé avec agrégation des quantiles

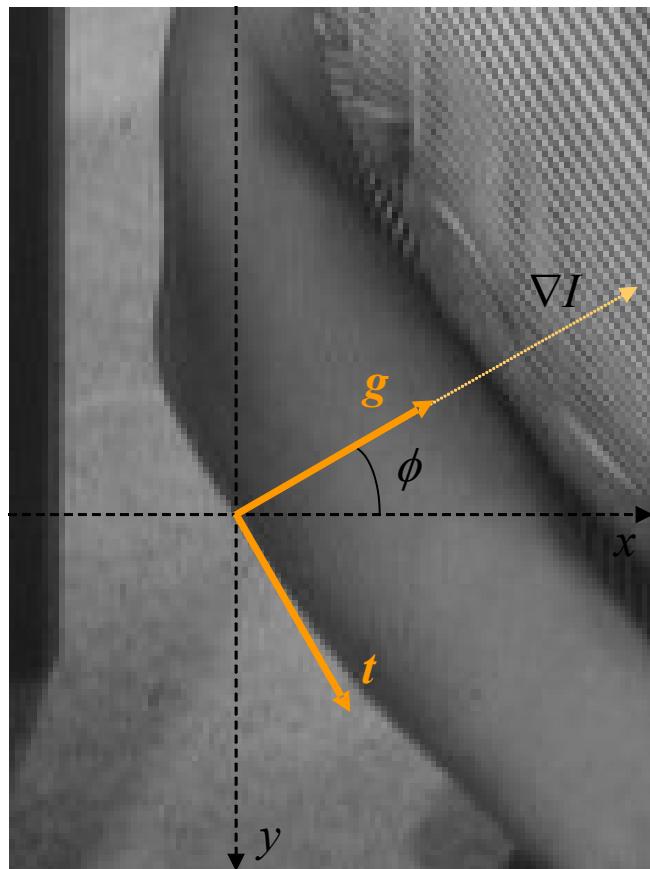


Image segmentée

Le modèle différentiel

Dans le modèle différentiel, on considère l'image comme une fonction continue $f(x,y)$, dont on étudie le comportement local à l'aide de ses dérivées.

Une telle étude, fondée sur la formule de Taylor, n'a de sens que si la fonction f a une certaine régularité, ce qui constitue le problème clef des méthodes différentielles.



Au premier ordre, on peut ainsi associer à chaque point (x,y) un repère propre (t,g) , où le vecteur t donne la direction de l'*isophote* (ligne de variation minimale) et g la direction orthogonale, celle du *gradient*.

Grâce au plongement dans le continu, le modèle différentiel permet en outre d'exprimer un grand nombre d'opérations d'analyse en termes d'*équations aux dérivées partielles* (EDP), ce qui permet de donner un fondement mathématique satisfaisant aux traitements et aussi de fournir des méthodes pour les calculer, par des schémas numériques de résolution.

Estimation des dérivées partielles

Les approximations les plus simples des dérivées directionnelles se font par différences finies calculées par convolution avec des noyaux très simples :

Par ex : $[-1 \ 1]$, pour l'approximation de $\frac{\partial f}{\partial x}$, et : $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$, pour l'approximation de $\frac{\partial f}{\partial y}$

On utilise plus souvent $[-1 \ 0 \ 1]$, respectivement $\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ qui produisent des frontières plus épaisses, mais bien centrées (phase nulle).

Ces opérations étant très sensibles au bruit, on les combine en général avec un filtre lisseur dans la direction orthogonale à celle de dérivation, par ex par le noyau suivant (ou sa transposée) : $[1 \ 2 \ 1]$

Le calcul des dérivées directionnelles en x et en y revient finalement à la convolution avec les noyaux suivants, respectivement :

$$f_x[i, j] = (f * h_x)[i, j] \quad , \text{ avec : } \quad h_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad h_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad (\text{Masques de Sobel})$$
$$f_y[i, j] = (f * h_y)[i, j]$$

On peut ensuite calculer la norme du gradient :

$$\|\nabla f[i, j]\|_2 = \sqrt{f_x[i, j]^2 + f_y[i, j]^2}$$
$$\|\nabla f[i, j]\|_1 = |f_x[i, j]| + |f_y[i, j]|$$
$$\|\nabla f[i, j]\|_\infty = \max\{|f_x[i, j]|, |f_y[i, j]|\}$$

Et son orientation :

$$\arg(\nabla f[i, j]) = \arctan\left(\frac{f_y[i, j]}{f_x[i, j]}\right)$$

Estimation des dérivées partielles



Original



Noyau $[-1 \ 1]$



Noyau $[-1 \ 0 \ 1]$

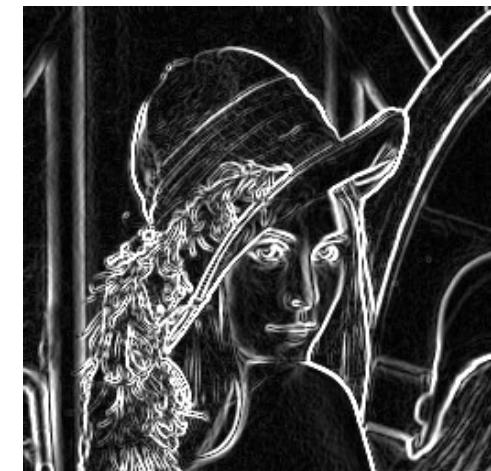
Notons que les noyaux dérivateurs sont à somme nulle, la transformée de Fourier de leur réponse impulsionnelle est donc nulle à l'origine.



Gradient horizontal (Sobel)



Gradient vertical (Sobel)



Module du gradient de Sobel

Estimation des dérivées partielles

L'approximation par différences finies la plus simple de la dérivée seconde est la convolution par le noyau :

$$[1 - 2 1] \text{ , pour l'approximation de } \frac{\partial^2 f}{\partial x^2} \text{, et : } \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \text{ , pour l'approximation de } \frac{\partial^2 f}{\partial y^2}$$

Le laplacien $\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$ peut donc être approximé par l'opérateur linéaire suivant :

$$\begin{bmatrix} 1 & & \\ & -4 & \\ 1 & & \end{bmatrix} \quad \textit{Laplacien en 4-connexité}$$

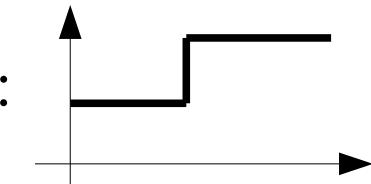
$$\text{, ou encore} \quad \begin{bmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad \textit{Laplacien en 8-connexité}$$



Exercice : Unsharp Masking

- Expliquer pourquoi la convolution par le noyau $[1 -2 1]$ approxime $\frac{\partial^2 f}{\partial x^2}$
- Montrer son comportement sur une simple transition 1d :
- Expliquer pourquoi le noyau de convolution suivant augmente le contraste local (voir TP1) :

$$\begin{bmatrix} -1 \\ -1 & \mathbf{5} & -1 \\ -1 \end{bmatrix}$$



f

$$f * \begin{bmatrix} -1 \\ -1 & \mathbf{5} & -1 \\ -1 \end{bmatrix}$$

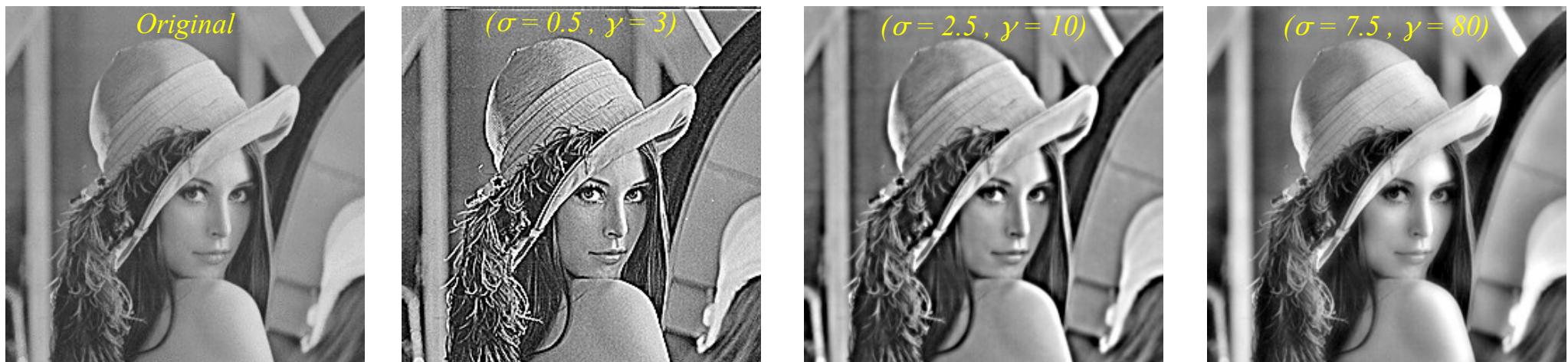
Limite des différences finies

Les opérateurs de traitement d'images possèdent une certaine portée correspondant à l'amplitude du voisinage qui interagit.

Un exemple, l'opération de réhaussement de contraste (unsharp masking) :

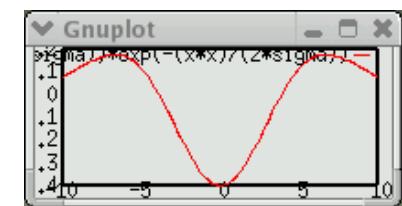
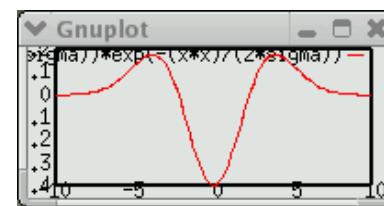
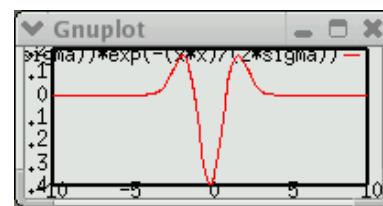
$$R_f[x, y] = f[x, y] - \gamma \cdot \Delta f[x, y] \quad (\gamma: \text{gain})$$

En soustrayant le laplacien à l'image original, on augmente le contraste.
Mais le contraste est une notion multi-échelle :



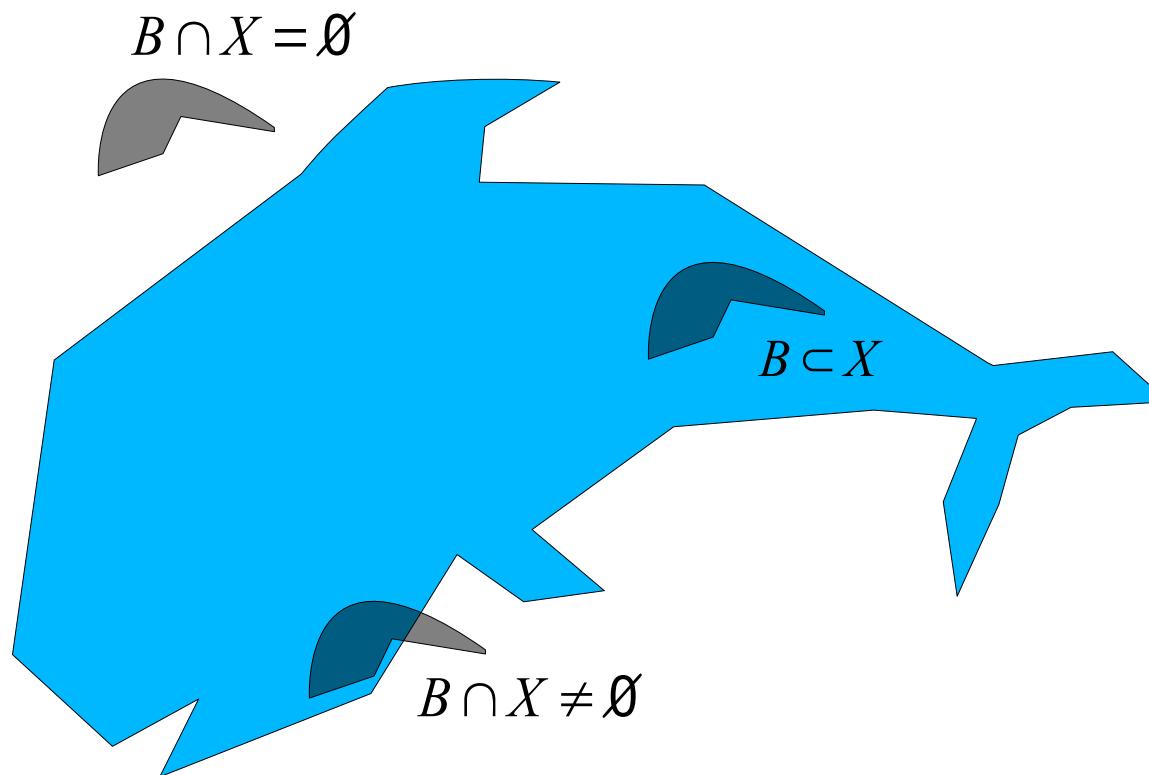
Il faut donc des mécanismes pour adapter la portée des opérateurs :

Laplaciens calculés par dérivées secondes de noyaux gaussiens, pour $\sigma = 1$, puis 5 , puis 15 :



Le modèle ensembliste

En *morphologie mathématique*, l'image est considérée comme un *ensemble*, dont on étudie les propriétés en fonction de relations locales avec un ensemble de référence (élément structurant) en termes d'*intersection* et d'*inclusion* (relations en tout-ou-rien).



Érosion et Dilatation

Les transformations morphologiques sont définies à partir des 2 opérateurs ensemblistes de base que sont l'*érosion* et la *dilatation*



Original (Matisse - 1952)



$$\varepsilon_B(X) = \{x \in \mathbb{R}^2 ; B_x \subset X\}$$



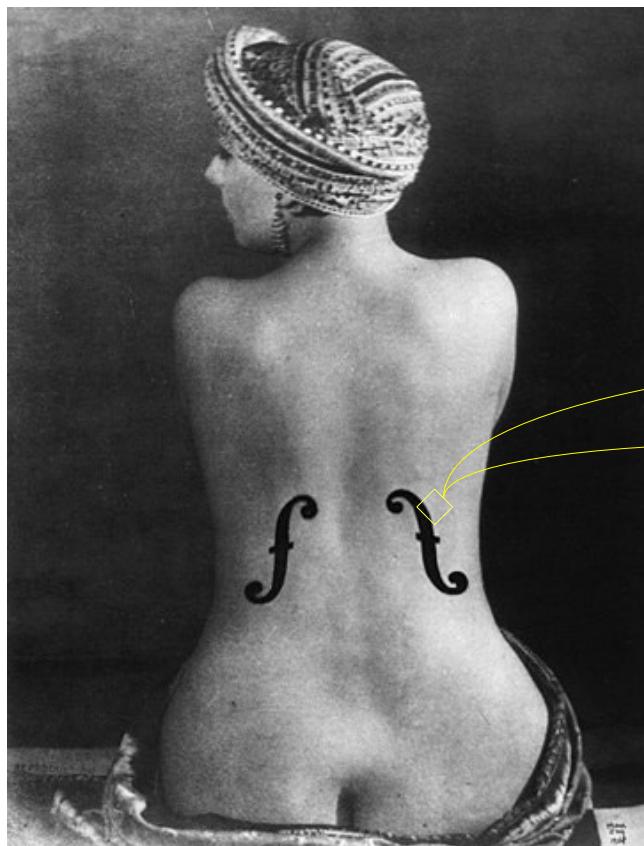
$$\delta_B(X) = \{x \in \mathbb{R}^2 ; B_x \cap X \neq \emptyset\}$$

(élément structurant : disque)

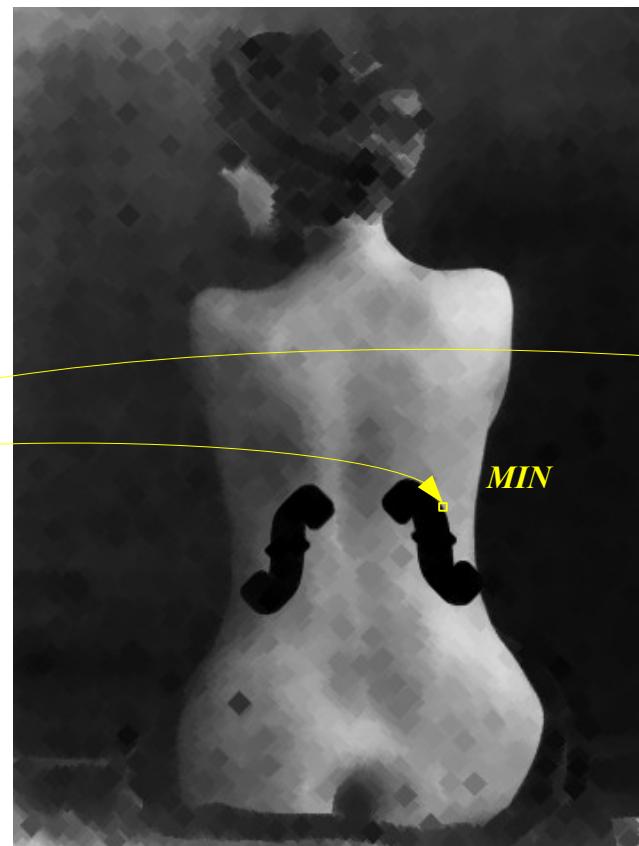
Érosion et Dilatation

L'érosion et la dilatation, et par suite l'ensemble des transformations morphologiques, se généralisent des ensembles (Images binaires) aux fonctions (Images en niveaux de gris) par l'intermédiaire des ensembles de niveaux :

$$I_n = \{x \in \mathbb{R}^2 ; I(x) \leq n\}$$

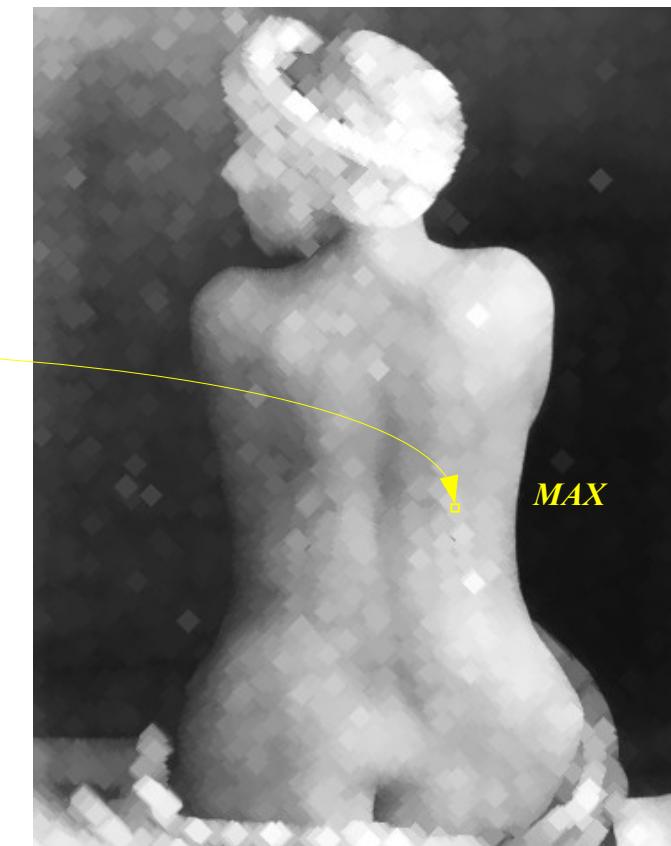


Original (Man Ray - 1924)



$$\varepsilon_B(I)$$

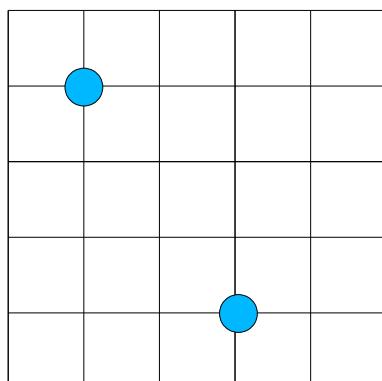
(élément structurant : losange)



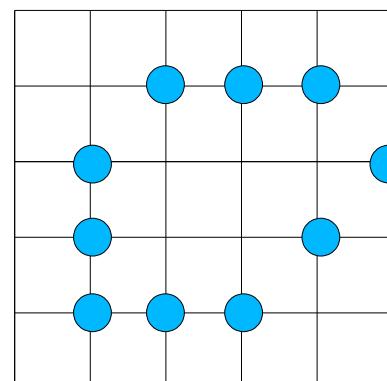
$$\delta_B(I)$$

Le modèle discret

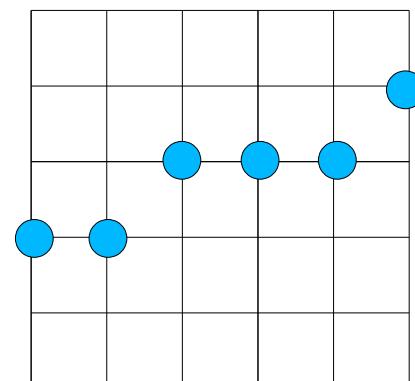
La *géométrie discrète* est une discipline au moins aussi ancienne que le traitement d'images. Alors que le modèle différentiel considère les structures géométriques (courbes, surfaces, droites, etc) comme des approximations numériques de leurs homologues continues, ou que le modèle fréquentiel traduit la discréttisation en termes de perte d'information, le modèle discret, lui, intègre l'espace échantillonné comme cadre mathématique, et s'efforce de donner un cadre formel aux structures géométriques discrètes : définition, propriétés, théorèmes,...



Quelle est la distance entre les 2 points ?



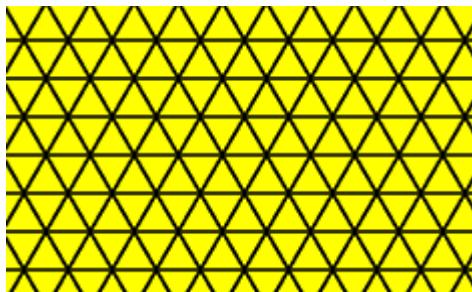
Qu'est-ce qu'un trou ?



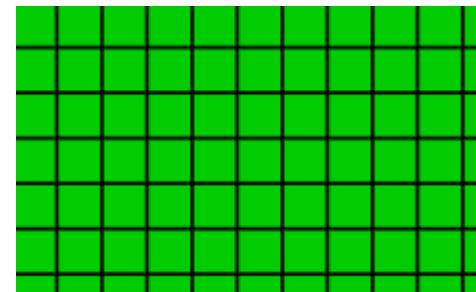
Qu'est-ce qu'une droite ?

Pavages du plan

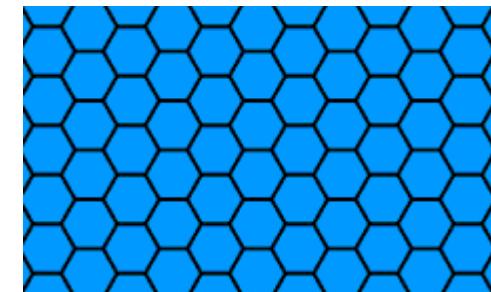
Un pavage du plan est une partition du plan en cellules élémentaires (pixels).
Il n'existe que 3 pavages réguliers du plan :



triangulaire

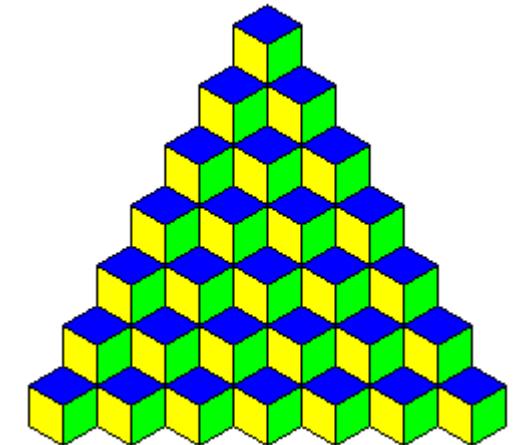
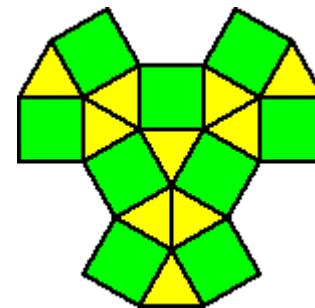
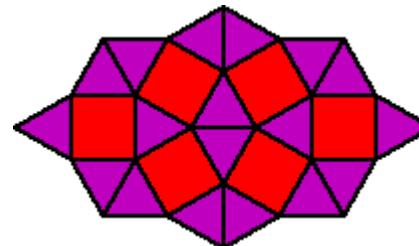
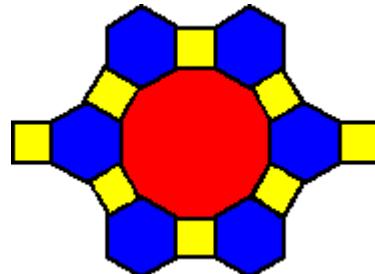


carré



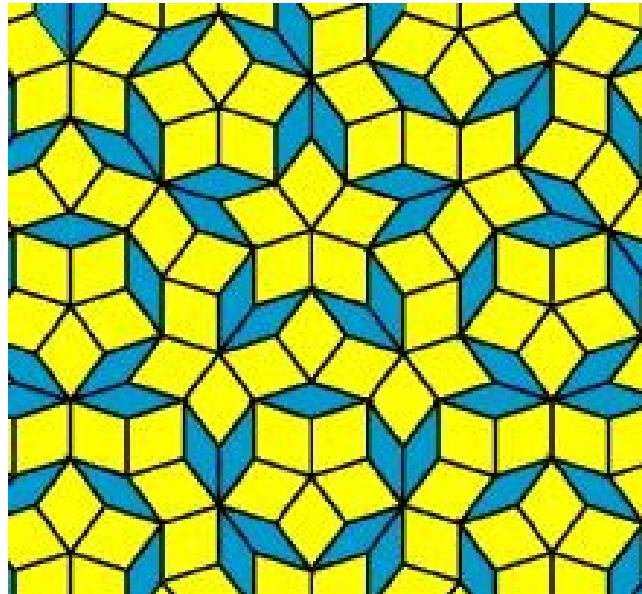
hexagonal

... mais de nombreux pavages irréguliers :

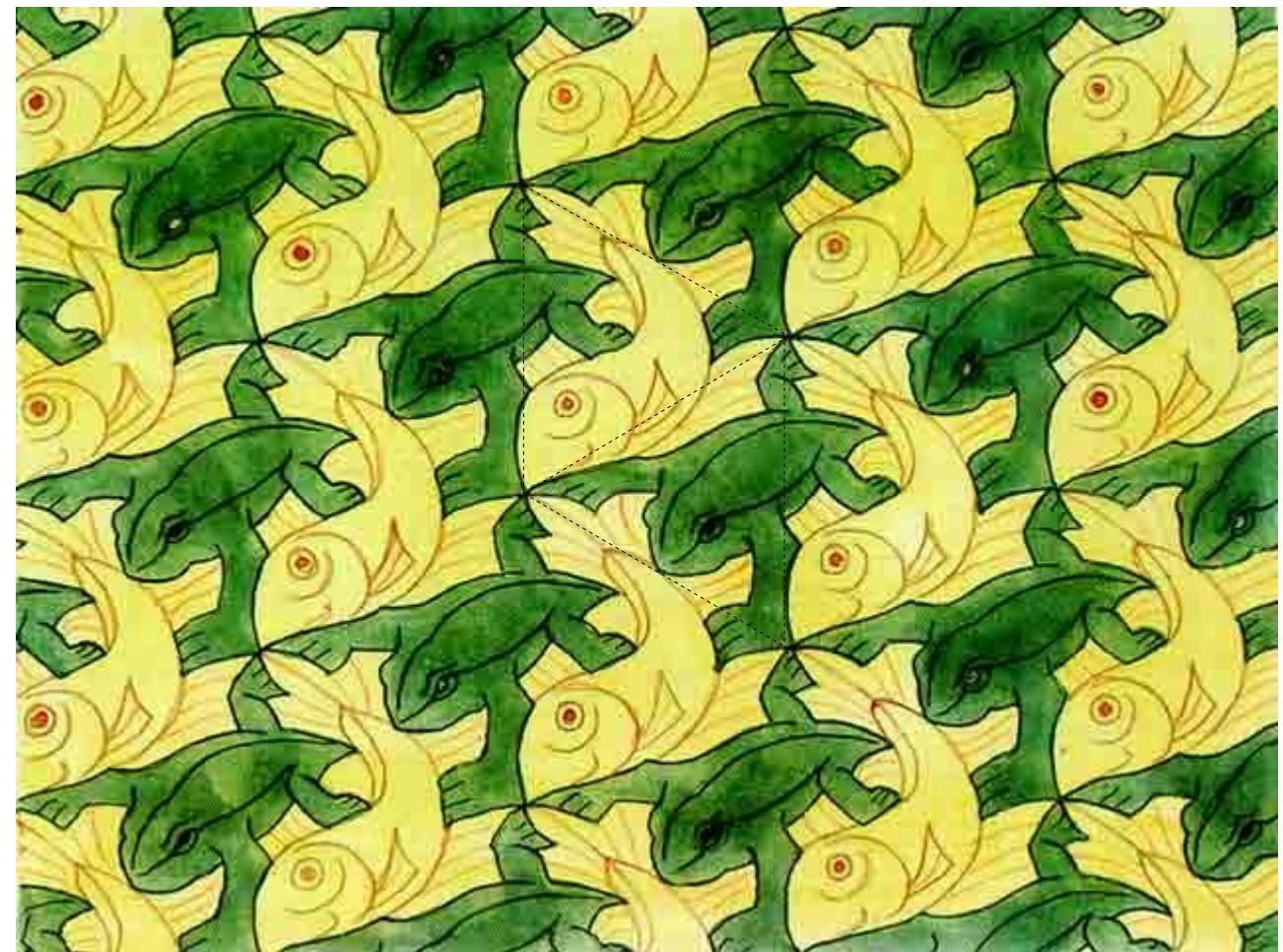


Pavages du plan

D'autres pavages irréguliers du plan...



Pavage apériodique de Penrose



Pavage périodique d'Escher

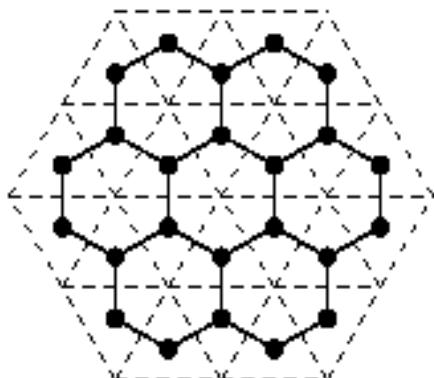
Pavages et maillages

A tout pavage du plan on peut associer un *graphe* où les *sommets* (nœuds) représentent les cellules élémentaires, et où les *arêtes* représentent la relation d'*adjacence* entre les cellules (2 cellules sont adjacentes si elles ont un côté en commun). Un tel graphe est un *maillage* du plan.

Les pavages et les maillages réguliers sont *duaux* :

Questions :

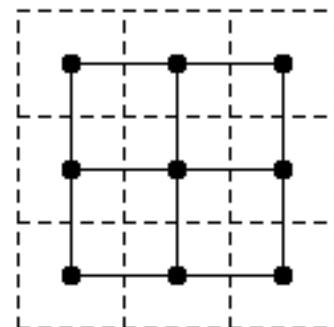
- représentation dans \mathbb{Z}^2 ?
- combien de directions ?
- récursivité ?



Pavage triangulaire



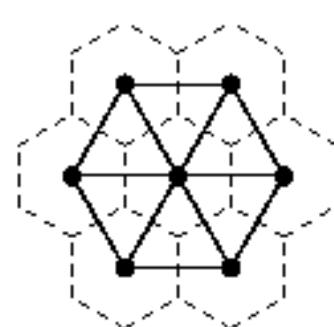
Maillage hexagonal



Pavage carré



Maillage carré



Pavage hexagonal



Maillage triangulaire

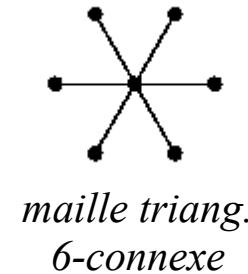
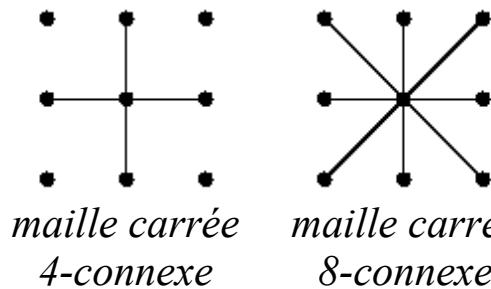
Maillage et connexité

Les relations topologiques dans les images discrètes sont définies à partir de la relation de connexité induite par le graphe du maillage (X, S) , où X représente les sommets et S les arêtes.

$$X \subset \mathbb{Z}^2; S \subset X^2$$

Soient x et y 2 points de X , par définition x et y sont *voisins* si :

$$x \approx y \Leftrightarrow (x, y) \in S$$



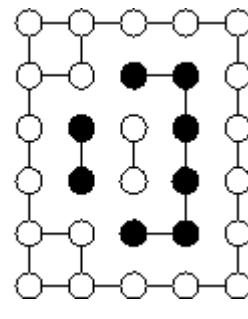
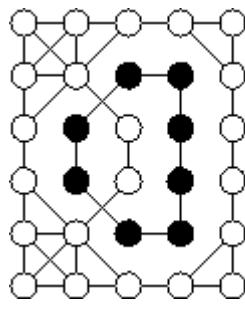
La clôture transitive de la relation de voisinage est une relation d'équivalence
« il existe un chemin connexe entre x et y » :

$$x \sim y \Leftrightarrow \exists \{x_1, \dots, x_n\} / x \approx x_1, \dots, x_i \approx x_{i+1}, \dots, x_n \approx y$$

Les classes d'équivalence de cette relation s'appellent les *composantes connexes* de X

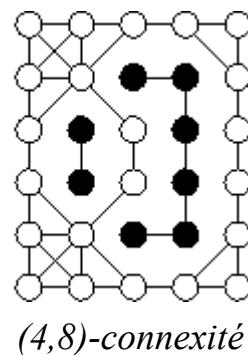
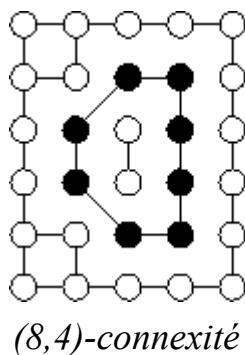
Topologie dans la maille carrée

Dans la maille carrée, la notion de *trou* dans un objet X ($X \subset \mathbb{Z}^2$), qui doit correspondre à une composante connexe finie du complémentaire X^c , n'est pas bien définie...



Ce problème est lié à la validité du *théorème de Jordan*, selon lequel une courbe simple fermée sépare le plan en 2 composantes connexes, dont une bornée.

...sauf si l'on considère des *connexités différentes* pour X et pour X^c :



Le *théorème de Jordan* est vérifié pour ces connexités.



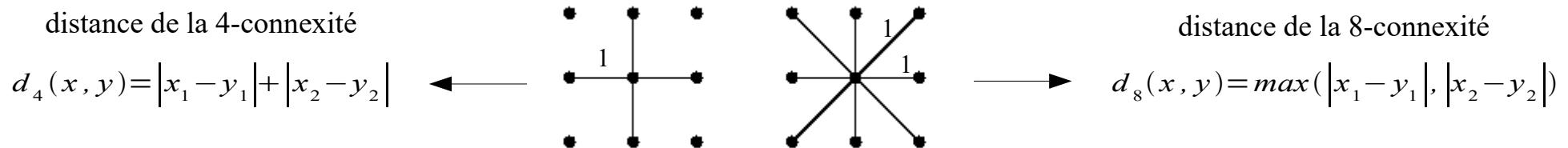
Questions :

combien de composantes connexes, combien de trous compte l'image ci-contre

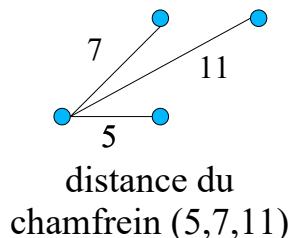
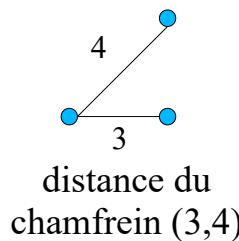
- en (8,4)-connexité ?
- en (4,8)-connexité ?

Métrique dans la maille carrée

Le graphe du maillage induit également une distance dans le plan discret, la distance entre 2 points x et y étant définie par la longueur du plus court chemin connecté entre x et y . En pondérant toutes les arêtes du maillage par la valeur 1, on trouve :

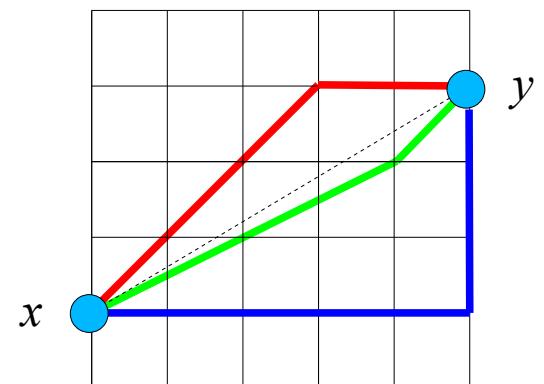


On peut aussi pondérer différemment les arêtes du maillage 8-connexe, voire utiliser des maillages plus complexes (i.e. des voisinages plus grands) :

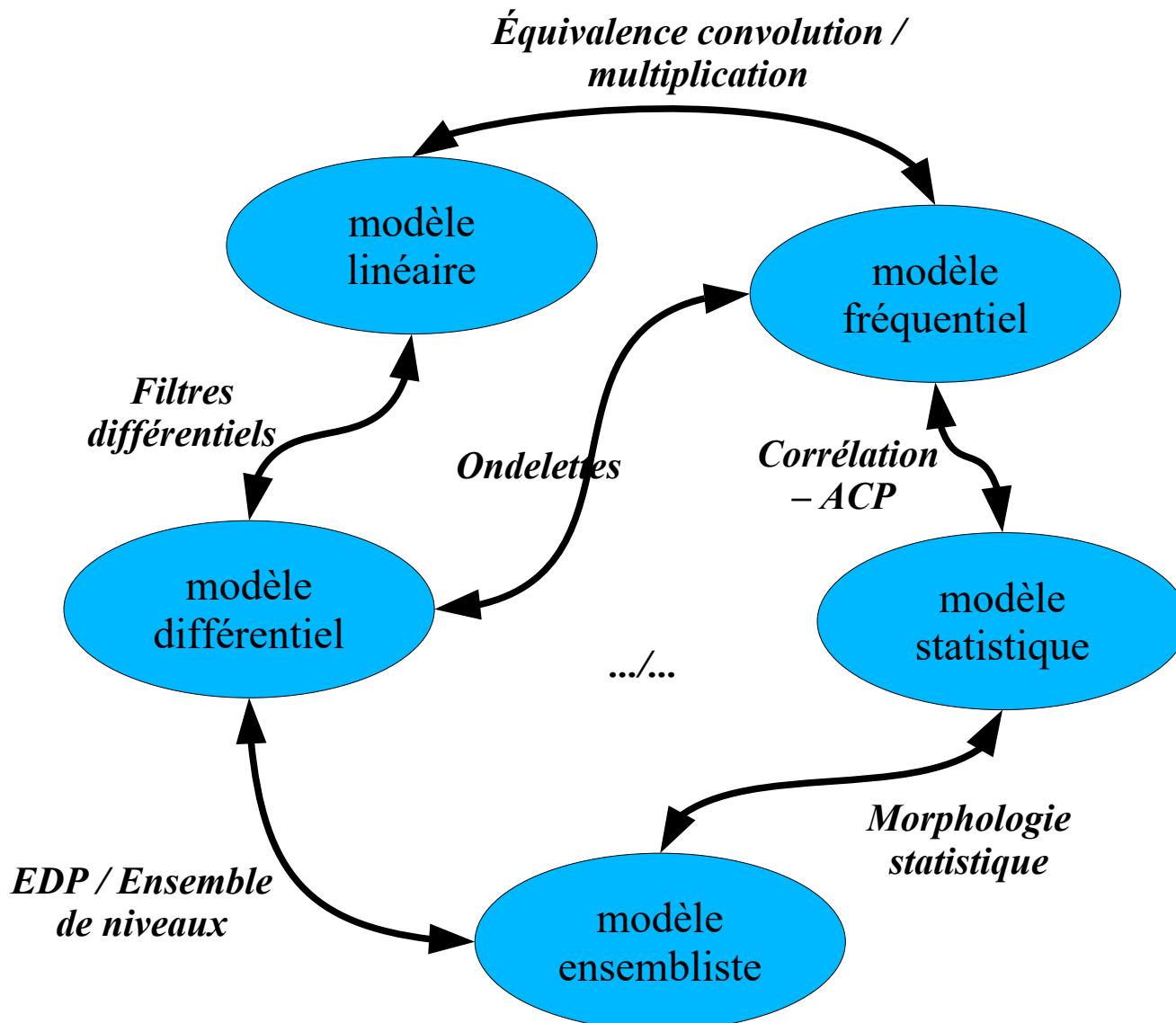


Questions :

calculer les distances $d_4(x, y)$, $d_8(x, y)$, $d_{ch(3,4)}(x, y)$, $d_{ch(5,7,11)}(x, y)$ entre les 2 points x et y ci-contre :



Conclusion



A retenir pour ce cours :

- image numérique
- échantillonnage
- quantification
- histogramme
- convolution
- représentation fréquentielle
- connexité et distance discrètes

Sources et bibliographie - Liens utiles

Livres :

- R.C. Gonzalez et Woods « *Digital Image Processing - 2d edition* » Addison Wesley 2002
- H. Maître (ss la direction de) « *Le traitement des images* » Hermes Lavoisier IC2 2003.
- S. Bres, J.M. Jolian, F. Lebourgeois « *Traitement et analyse des images numériques* » Hermes Lavoisier 2003
- M. Petrou, C. Petrou, « *Image Processing: the Fundamentals - 2nd edition* », Wiley, 2011 <https://the-eye.eu/public/Books/IT%20Various/Image%20Processing%C2%202nd%20Edition.pdf>
- D. Lingrand « *Introduction au Traitement d'images* » Vuibert 2004

Pages web :

- Univ. de Delft : <ftp://qiftp.tudelft.nl/DIPimage/docs/FIP2.3.pdf>
- CVOnLine - Univ. d'Édimbourg : <http://www.dai.ed.ac.uk/CVonline/>
- COURSERA : <https://www.coursera.org/learn/image-processing>