

TERI : Traitement et reconnaissance d'images

Cours Master IAD - 2006/07

Isabelle Bloch - ENST / Département Signal & Images

Henri Maître - ENST / Département Signal & Images

Florence Tupin - ENST / Département Signal & Images

Antoine Manzanera – ENSTA / Unité d'Électronique et d'Informatique



TERI – Objectifs du cours

L'objectif du cours « Traitement et Reconnaissance d'Images » est de proposer une *formation de base* en analyse d'images et reconnaissance des formes.

Il est destiné à tous les élèves suivant le master IA et Décision ; les liens entre l'Image, la Vision et l'Intelligence Artificielle seront abordés à plusieurs reprises.

On présentera les connaissances de base sur les images discrètes (*théorie de l'information*), leur représentation (*structures discrètes*), et leur exploitation (*filtrage et amélioration*).

L'analyse automatique des images sera développée à travers un problème phare : celui de la *segmentation* d'images.

On abordera également des notions de plus haut niveau sur la compréhension automatique du contenu des images (*classification et reconnaissance des formes*).

Le cours s'accompagnera d'une partie pratique : TP sur machine + TD exercices.

Enfin, une ouverture sur les *applications* et les *débouchés industriels* sera faite, qui préfigurera ce qui sera traité plus en détail dans les cours de filière (TDI).

TERI – Planning du cours 2006/07

- * 25/09 : *Introduction + échantillonnage* (4h) (Antoine Manzanera)
- * 2/10 : *Restauration et filtrage* (3h) (Antoine Manzanera)
- * 9/10 : *Introduction à la segmentation* (3h) (Henri Maître)
- * 16/10 : *Reconnaissance des formes et classification* (4h) (Isabelle Bloch)
- * 23/10 : *TP Introduction à l'image* (4h) (Isabelle Bloch) **ENST – Salle C128**
- * 30/10 : *TD* (4h) (Florence Tupin)
- * 6/11 : *Applications* (2h) *et présentation des stages* (2h) (F. Tupin, A. Manzanera)

Salle : 4600 – salle 101 sauf TP 23/10

Détails pratiques, Organisation des examens, Supports de cours,... :



<http://www.tsi.enst.fr/~bloch/TERI.html>

Introduction au TI – Plan du cours

I Développement du TI 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 Echantillonnage 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...

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,...

Historique du Traitement d'Images



Images de chambre à bulles

Restauration

Amélioration

EMPIRISME

VISION ACTIVE

1950

ABCDEFGHIJKLMNØP
QRSTUVWXYZAÆÏÖÜ
abcdefghijklmnp
qrstuvwxyzääáíóö&
1234567890(#+.?!?)

Caractères typographiés

Classification



RECONSTRUCTIONISME

Imagerie satellite et aérienne

MORPHOLOGIE MATHÉMATIQUE

Imagerie médicale



Télesurveillance et armement



EDP & SCALE SPACE

Reconstruction

Contrôle qualité

Détection

Poursuite

Robotique mobile

Localisation



Gestion des données multimedia

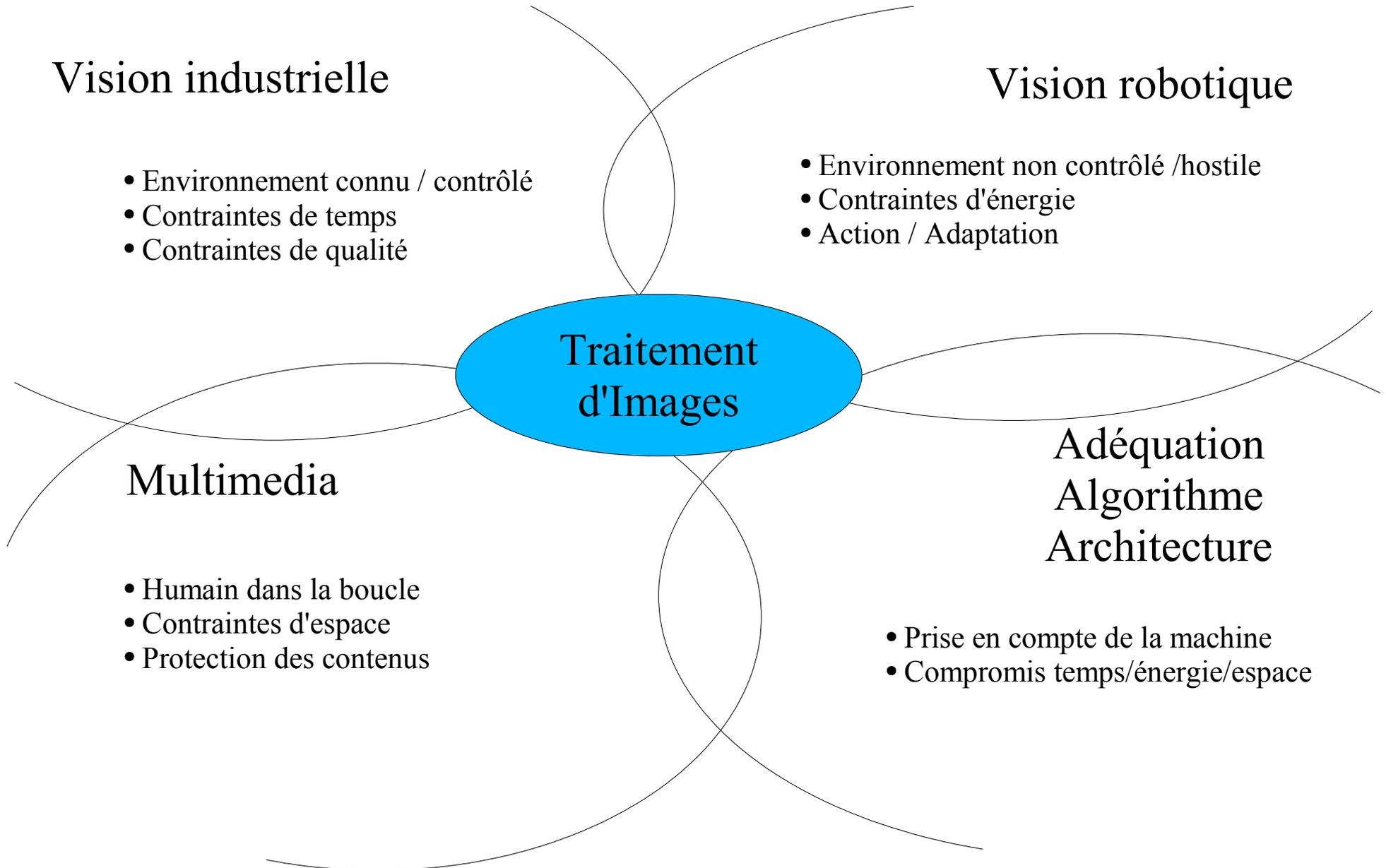
Indexation



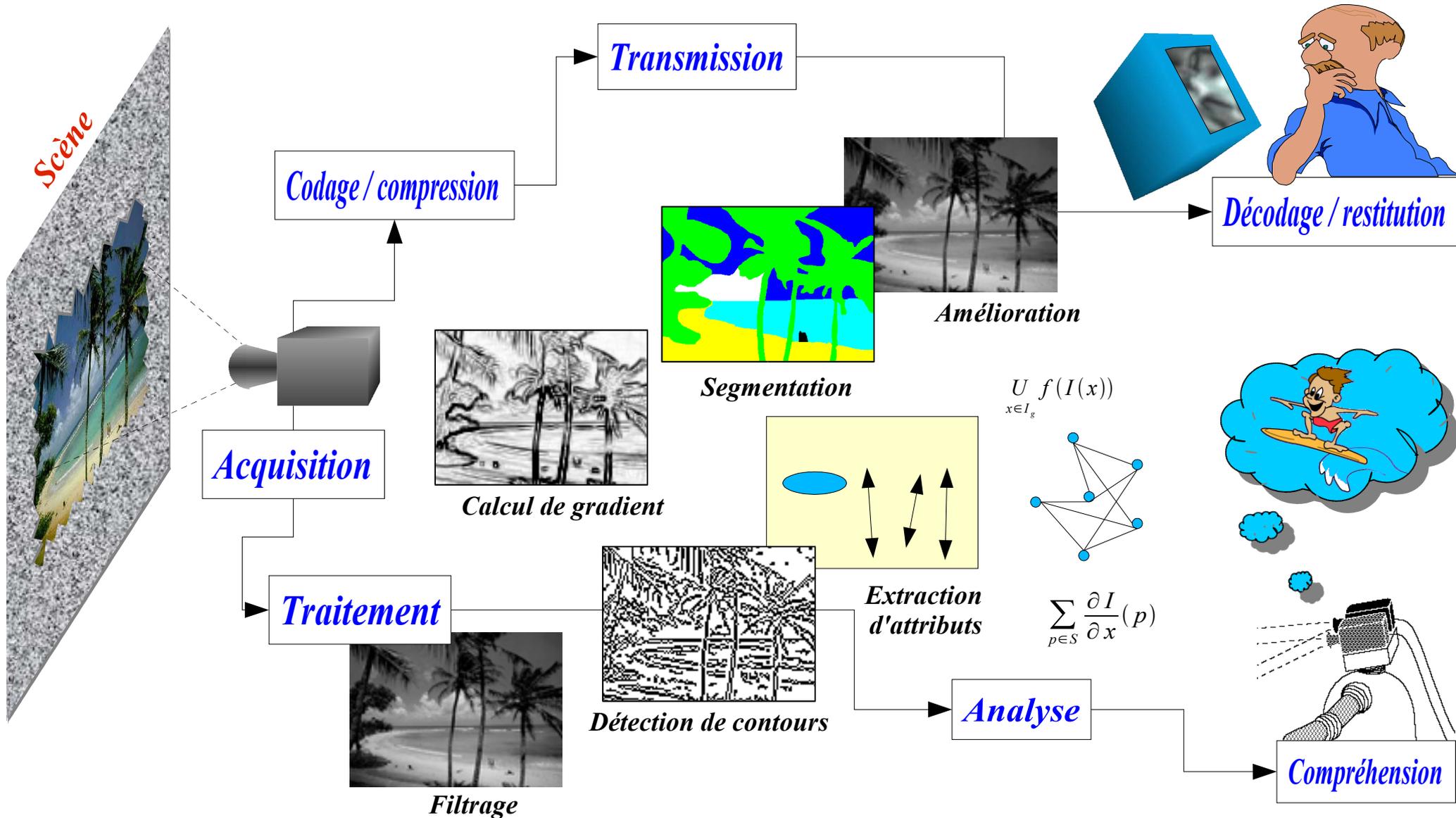
Compression

2005

TI & vision par ordinateur



Systemes à base de TI



TI & Intelligence Artificielle

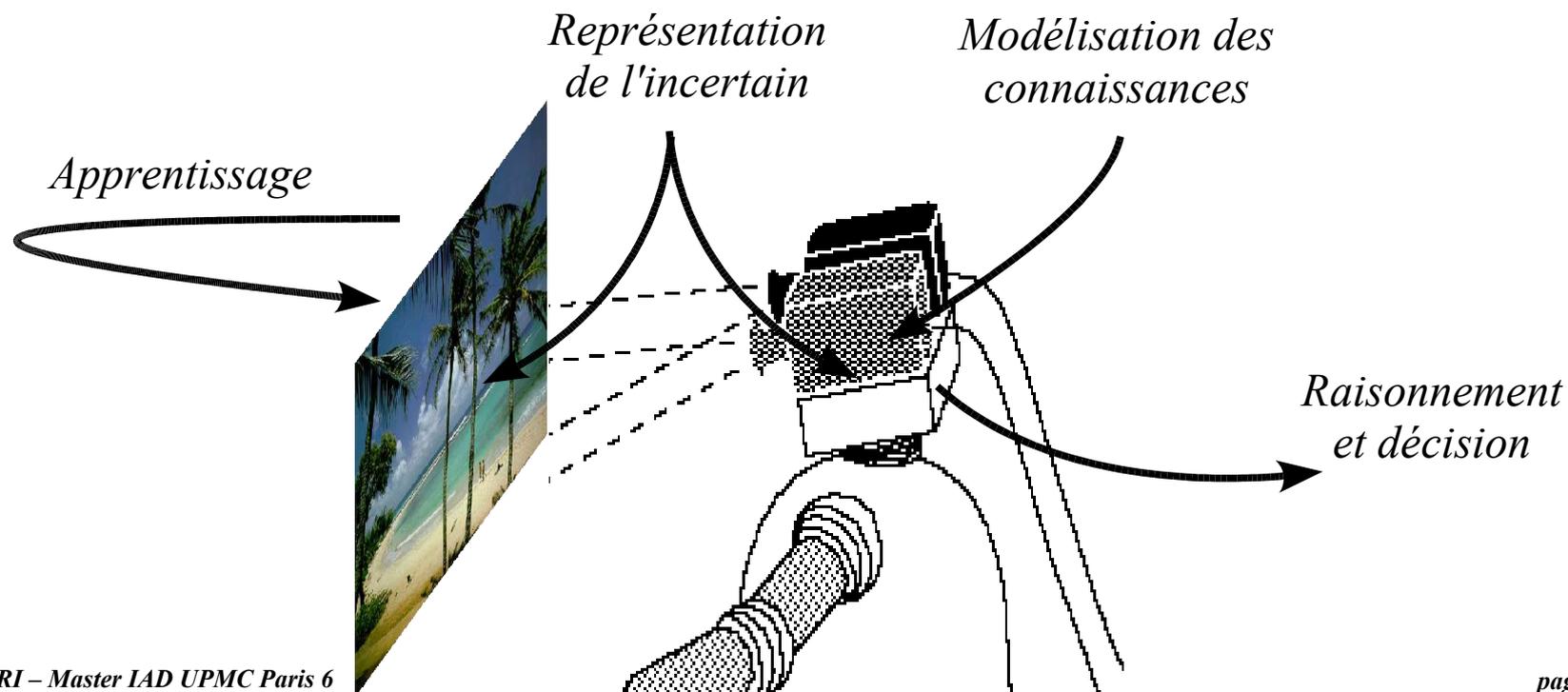
Dans la conception moderne de l'Intelligence Artificielle dite *située* (i.e. mise en situation) ou *incarnée* (i.e. introduite dans un « acteur »), la machine *agit* sur le monde extérieur, éventuellement *se déplace*, et aussi *perçoit* son environnement pour pouvoir *s'y adapter*.

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).



TI & Intelligence Artificielle

Réciproquement, le traitement d'image 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 la 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.
- Vision des batraciens.
- .../...

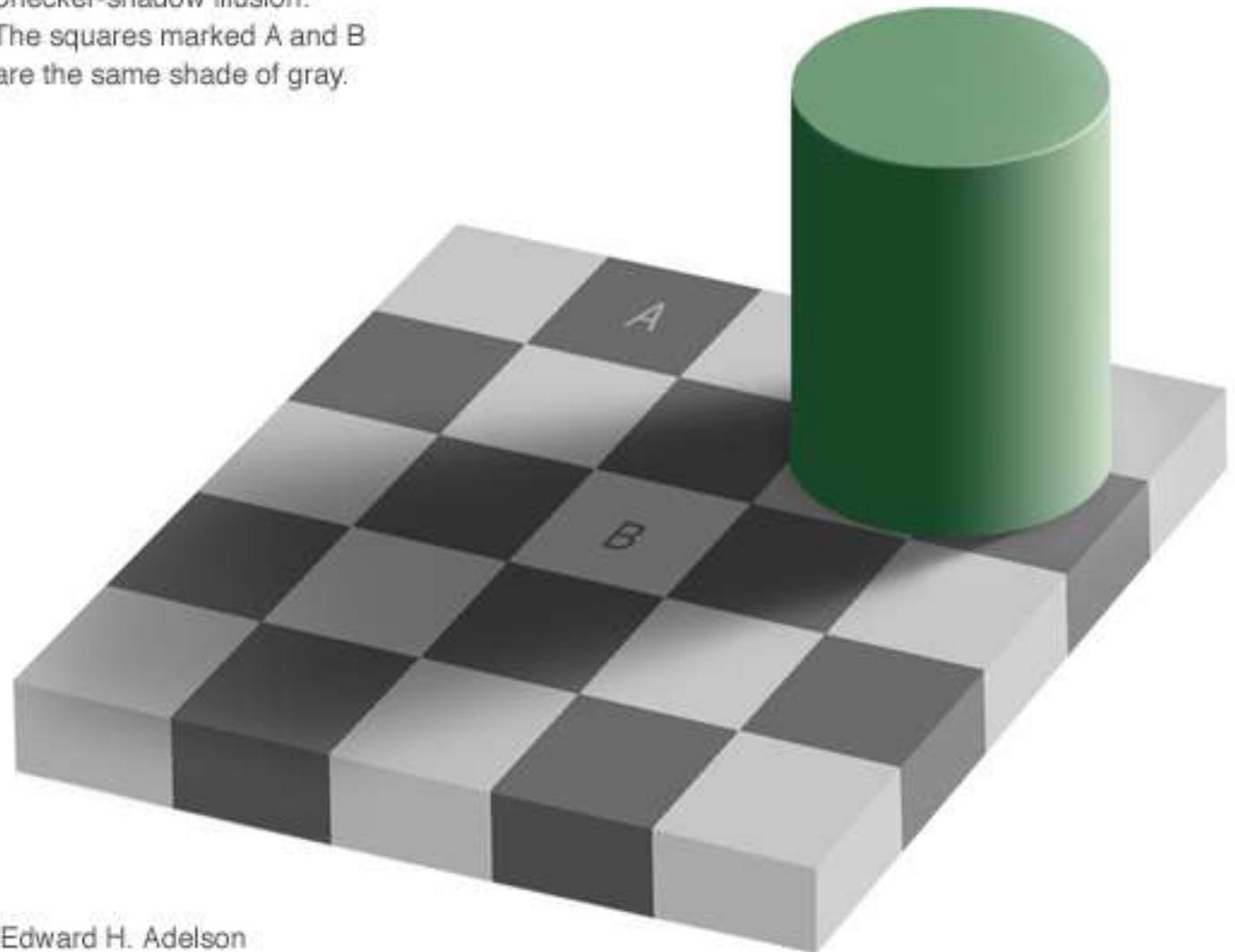
TI & perception visuelle

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)

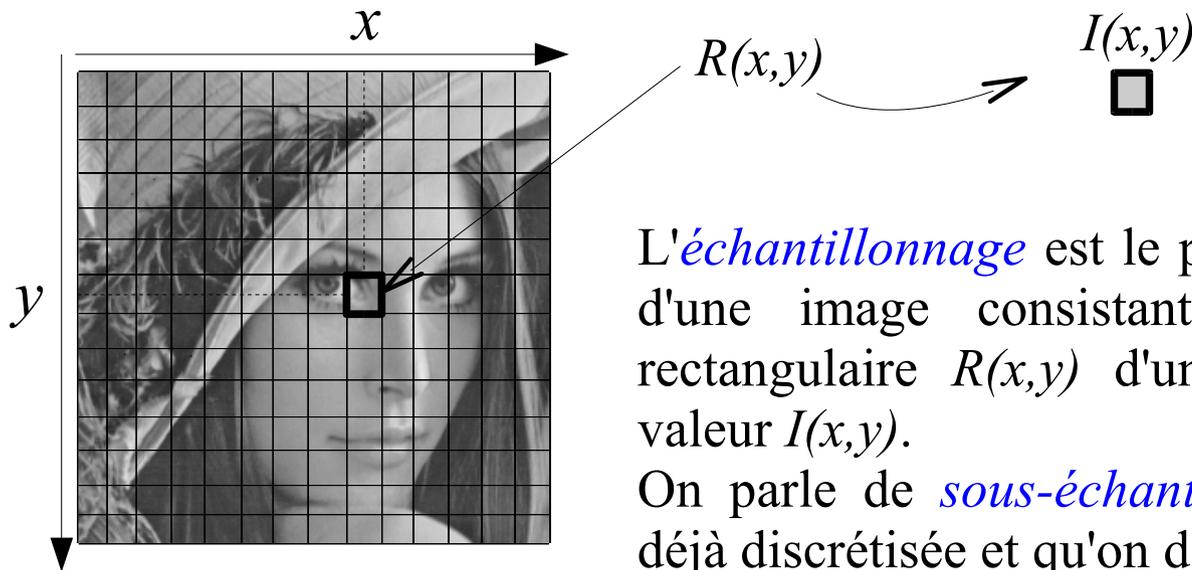


Edward H. Adelson

I-1 Types d'images, de capteurs

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

Images numériques



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

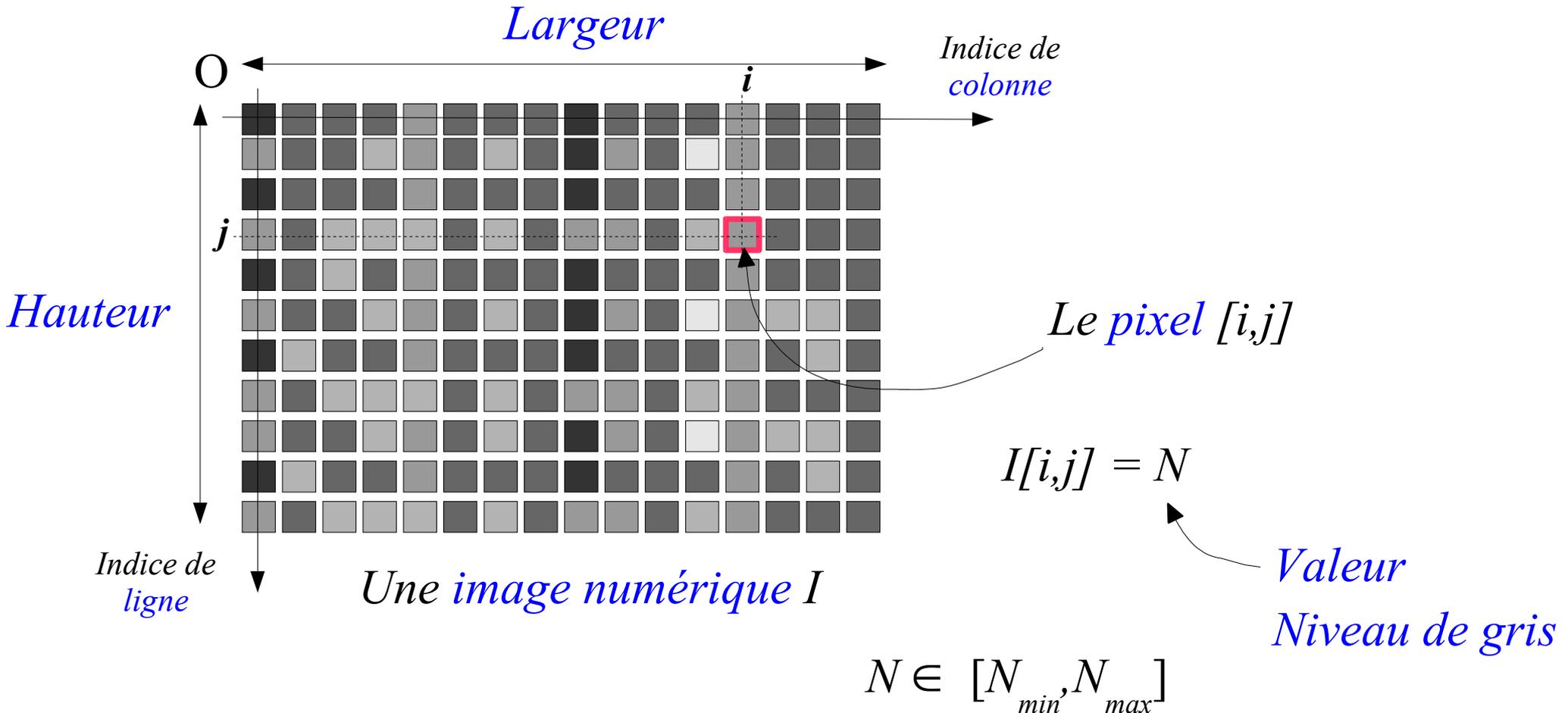
On parle de *sous-échantillonnage* lorsque l'image est déjà discrétisée et qu'on diminue le nombre d'échantillons.

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



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

$\text{Log}_2(N_{max} - N_{min}) = \text{dynamique}$

Echantillonnage et quantification

Résolution...

...spatiale :

Échantillonnage



256x256



128x128



64x64



32x32

...tonale :

Quantification



6 bits



4 bits



3 bits



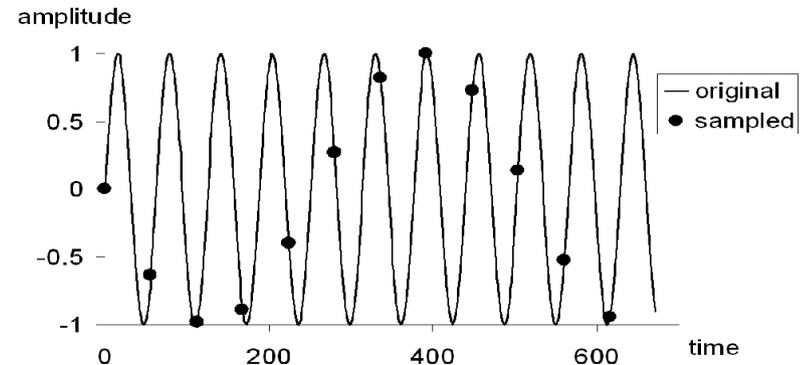
2 bits



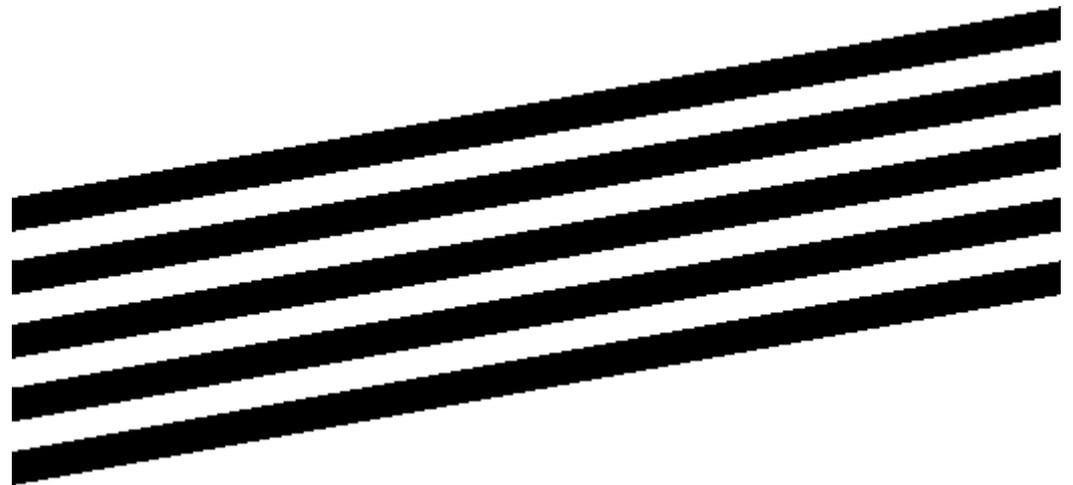
1 bit

Echantillonnage 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-contre, en 1d, le signal échantillonné « ressemble » à une sinusoïde de fréquence 8 fois plus faible :

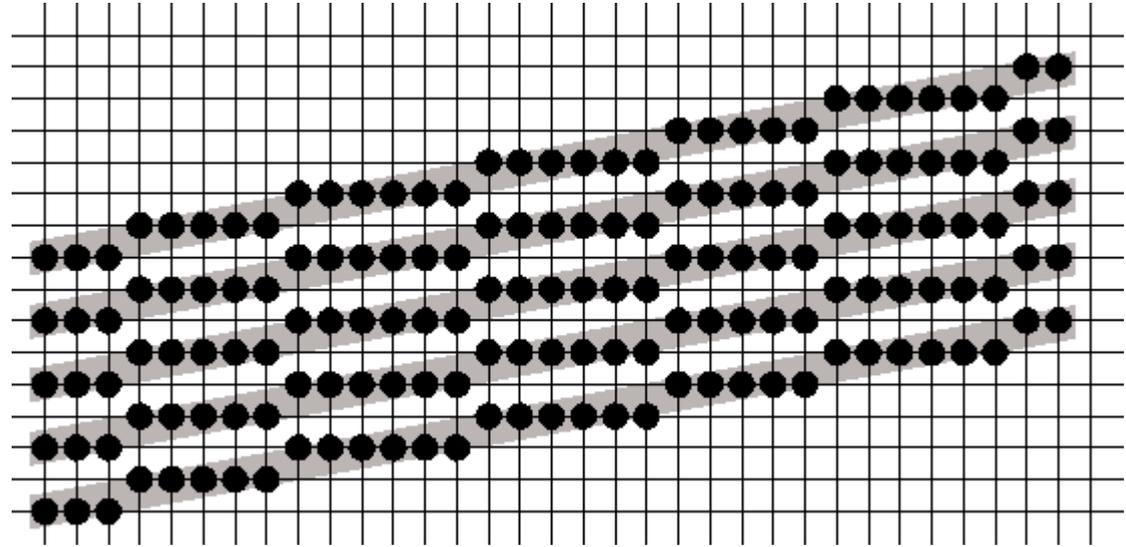


Ce phénomène appelé *aliasing* est encore pire en 2d, car il affecte *la fréquence et la direction* des structures périodiques. Imaginons par exemple qu'on souhaite échantillonner l'image correspondant aux bandes noires ci-contre :

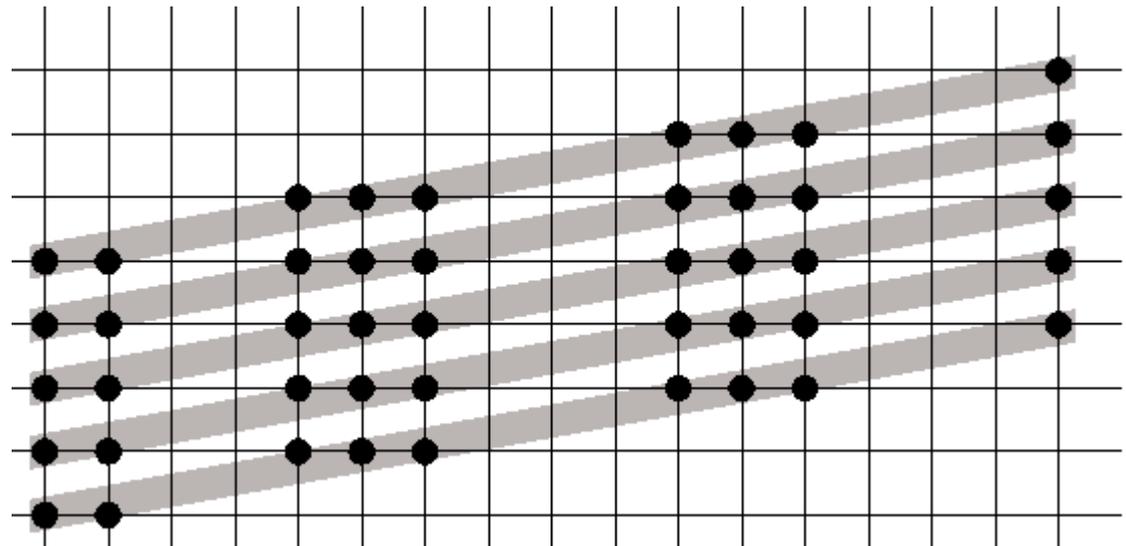


Echantillonnage et information

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 :



Echantillonnage et information

Un exemple, sur une image de synthèse :

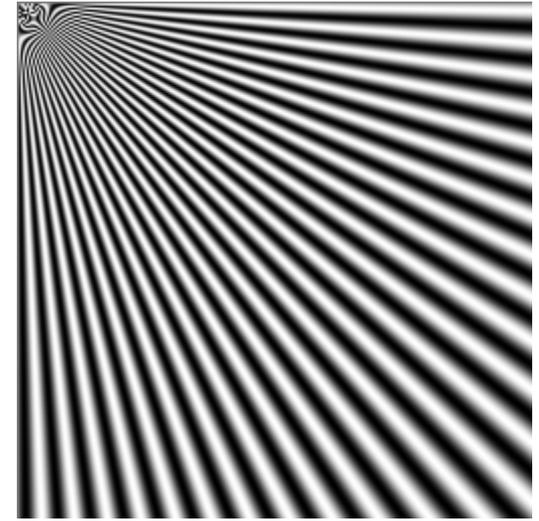


Image originale

Et sur une image naturelle :



Image sous-échantillonnée

Quantification et information

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



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.

III 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*...

Nous donnons dans la suite de ce cours une introduction aux outils les plus basiques ou les plus courants. Les autres seront traités ou approfondis dans la suite du cours TERI, ou dans le cours de filière TDI.

III-1 : 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

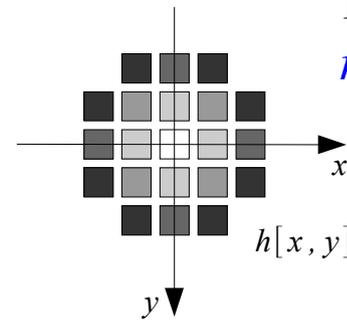
Soit I une image numérique.

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*



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

Propriétés de la convolution :

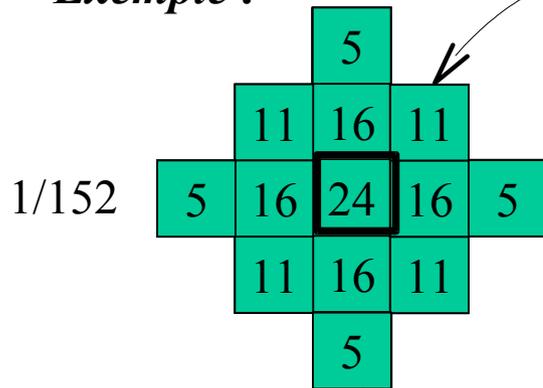
COMMUTATIVITÉ $h * g = g * h$

ASSOCIATIVITÉ $(h * g) * k = h * (g * k) = h * g * k$

DISTRIBUTIVITÉ / + $h * (g + k) = (h * g) + (h * k)$

La Convolution

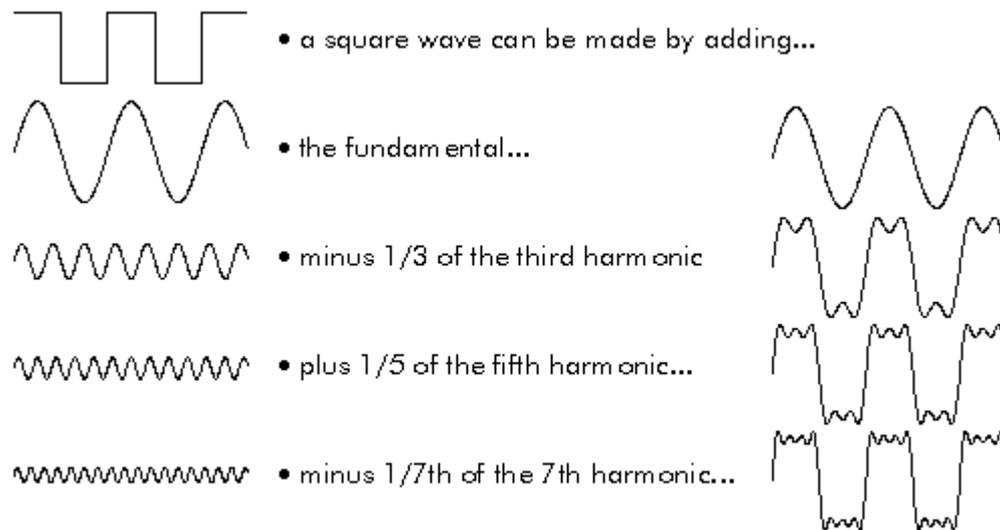
Exemple :



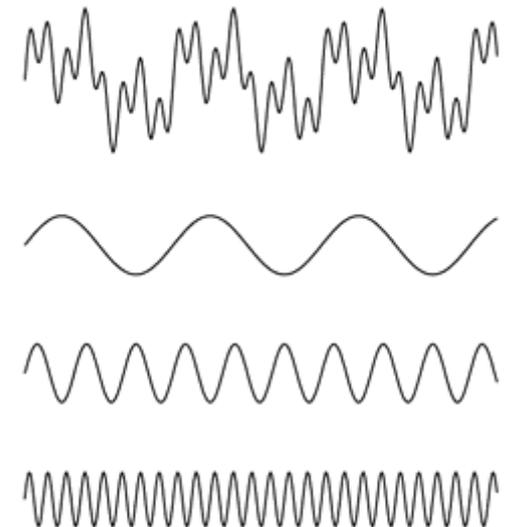
Pour calculer une convolution, on remplace la valeur de chaque pixel par la valeur du produit scalaire entre les valeurs du noyau de convolution et les valeurs du voisinage du pixel considéré (par rapport à l'origine (0,0) du noyau de convolution). Attention : implémentation « parallèle ».

III-2 : 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 sinusoides :



© **BORES** Signal Processing



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

La transformée de Fourier :

Outil fondamental d'analyse en traitement du signal, le pendant bidimensionnel de la TF et sa version discrète peut être appliqué 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 TDI espace d'échelles).

Transformée de Fourier

La transformée de Fourier permet la décomposition d'un signal f en *combinaison linéaire de sinusoides 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*

Directe :

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

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

Inverse :

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

Propriétés de la transformée de Fourier (1) :

ÉCRITURE SOUS FORME MODULE / PHASE

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

PÉRIODICITÉ

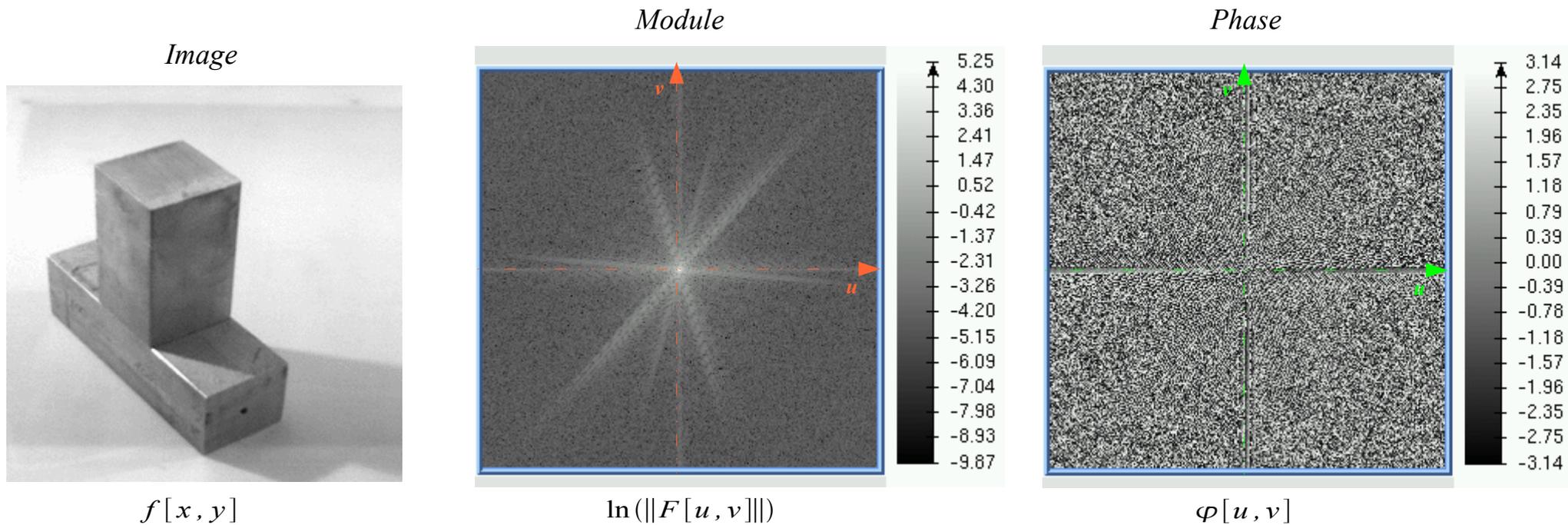
$$F[u, v] = F[u+w, v+h]$$

SYMÉTRIE

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

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

Transformée de Fourier



Propriétés de la transformée de Fourier (2) :

$$\text{si } \left\{ \begin{array}{l} f[x, y] \xrightarrow{\text{TF}} F[u, v] \\ f_1[x, y] \rightarrow F_1[u, v] \\ f_2[x, y] \rightarrow F_2[u, v] \end{array} \right\}$$

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

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

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

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

THÉORÈME DE PARSEVAL

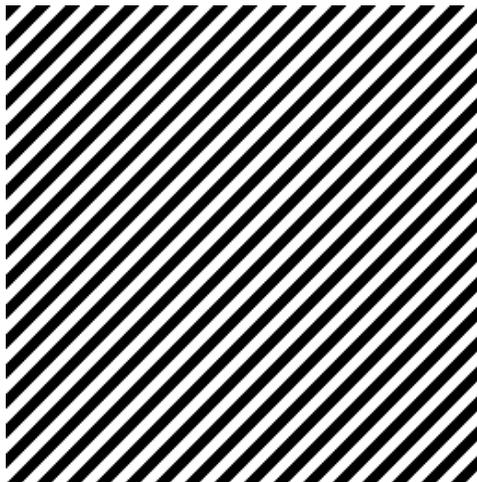
$$\sum_{x=0}^{w-1} \sum_{y=0}^{h-1} \|f[x, y]\|^2 = \frac{1}{wh} \sum_{u=0}^{w-1} \sum_{v=0}^{h-1} \|F[u, v]\|^2$$

QUIZZ - Transformée de Fourier

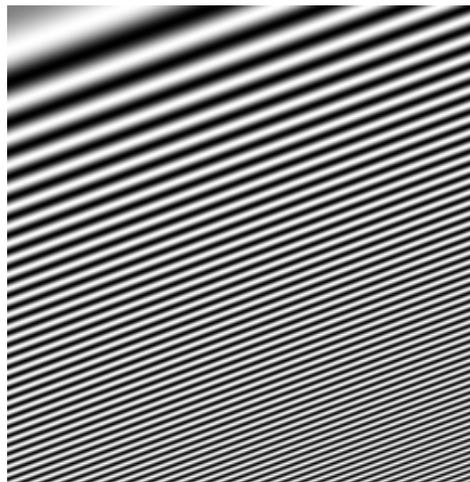
Attribuez à chaque image son spectre de Fourier



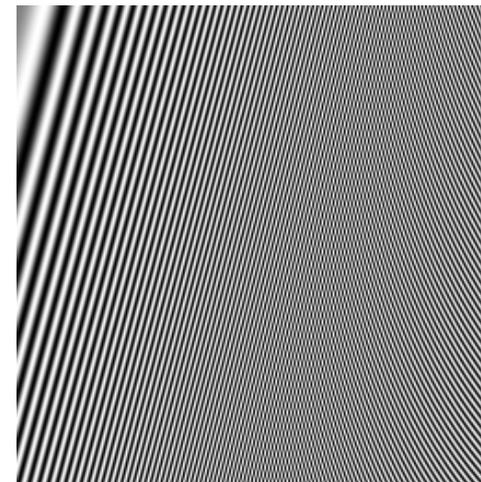
(1)



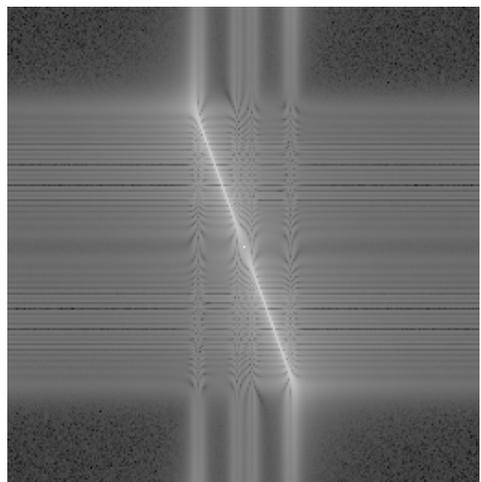
(2)



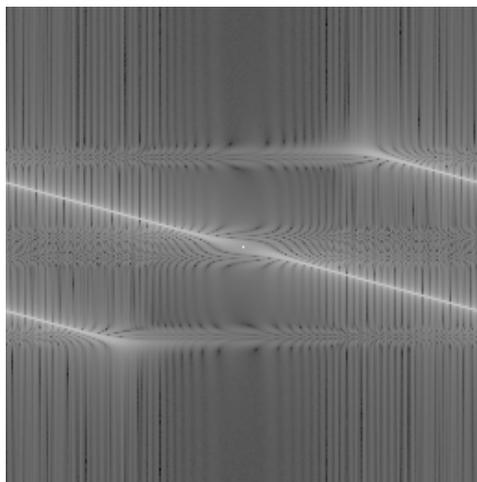
(3)



(4)



(a)



(b)



(c)



(d)

QUIZZ - Transformée de Fourier

Attribuez à chaque image son spectre de Fourier



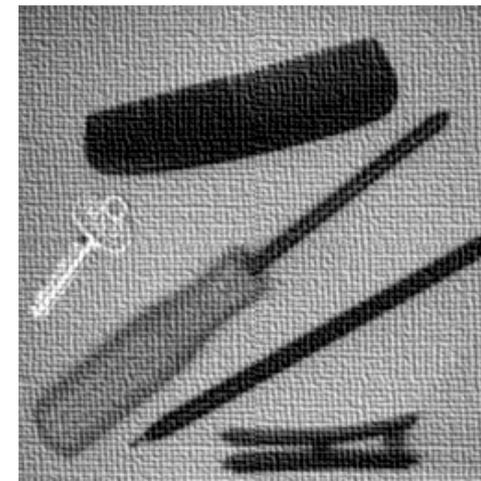
(1)



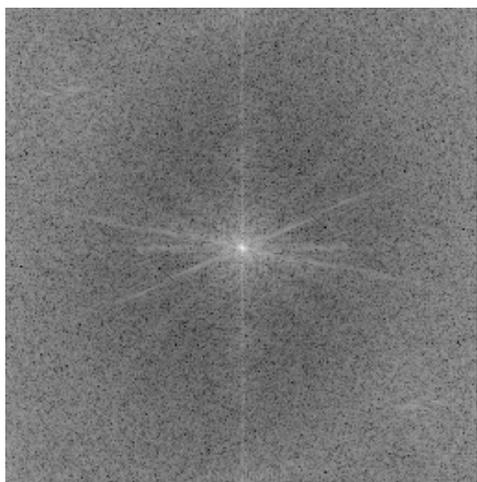
(2)



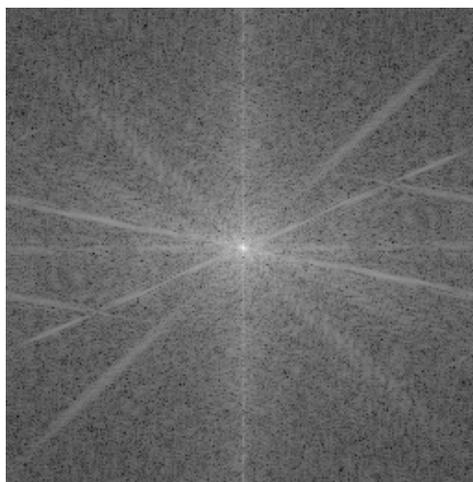
(3)



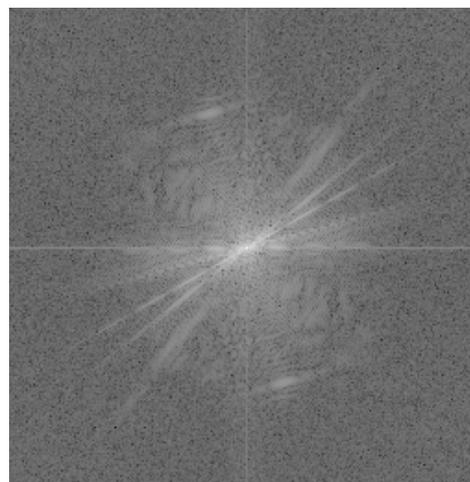
(4)



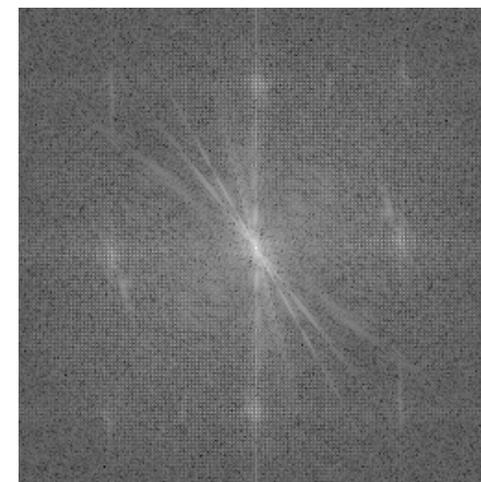
(a)



(b)



(c)



(d)

La norme de compression Jpeg utilise une représentation des images sous forme de *transformée en cosinus discrète bidimensionnelle (DCT2d)* :

Transformée directe :

$$C(u, v) = \frac{2}{n} \left(\frac{1}{\sqrt{2}} \right)^{\delta_{(0,0)}(u,v)} \sum_{x=0}^{n-1} \sum_{y=0}^{n-1} I(x, y) \cdot \cos \pi \frac{(2x+1)u}{n} \cdot \sin \pi \frac{(2y+1)v}{n}$$

Transformée inverse :

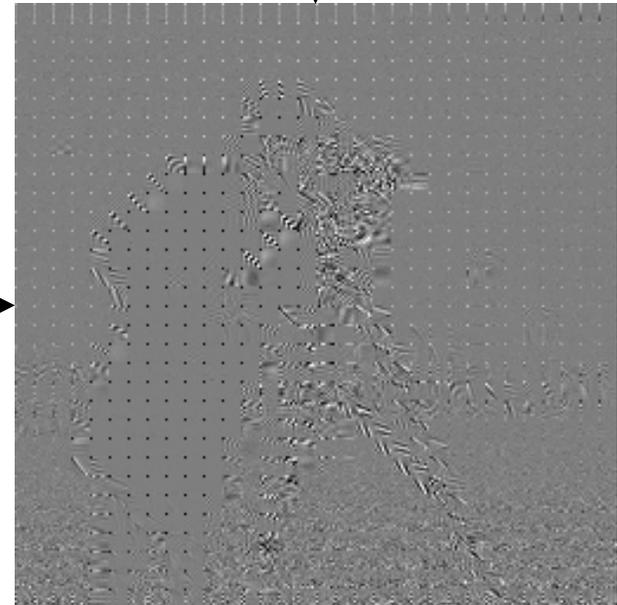
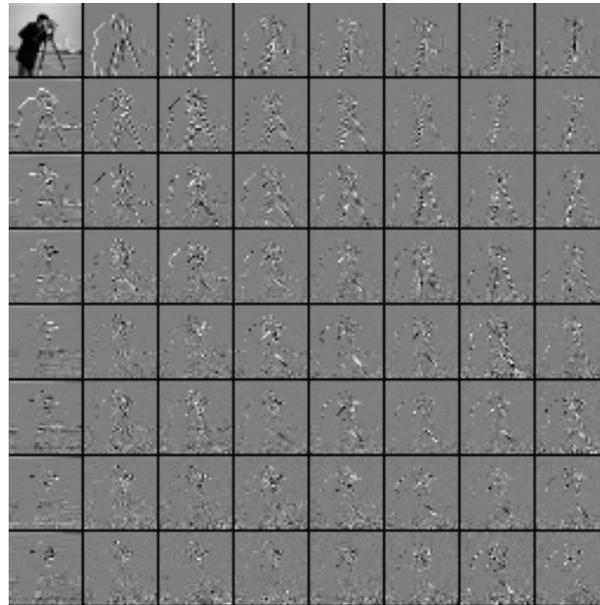
$$I(x, y) = \sum_{u=0}^{n-1} \sum_{v=0}^{n-1} \left(\frac{1}{\sqrt{2}} \right)^{\delta_{(0,0)}(u,v)} C(u, v) \cdot \cos \pi \frac{(2x+1)u}{n} \cdot \sin \pi \frac{(2y+1)v}{n}$$

Cette transformation est appliquée à *tous les blocs carrés d'un découpage régulier* de l'image, voir exemple ci-contre (n = 8) :

Le mode de représentation (1) de la DCT2d préserve la localisation des blocs de l'image originale : il illustre bien *l'irrégularité spatiale de la répartition de l'information*.

Le mode de représentation (2) de la DCT2d regroupe les pixels de tous les blocs par coefficient. Il permet de visualiser le *découpage en bande de fréquence*, et d'apprécier la différence de leur *pertinence* d'un point de vue *psychovisuel*.

La compression Jpeg exploite cette différence en *quantifiant plus fortement*, voire en supprimant les *fréquences les plus élevées* de la DCT2d



(2)

(1)

Transmission progressive d'image correspondant à une allocation de débit variable du codage DCT2d.

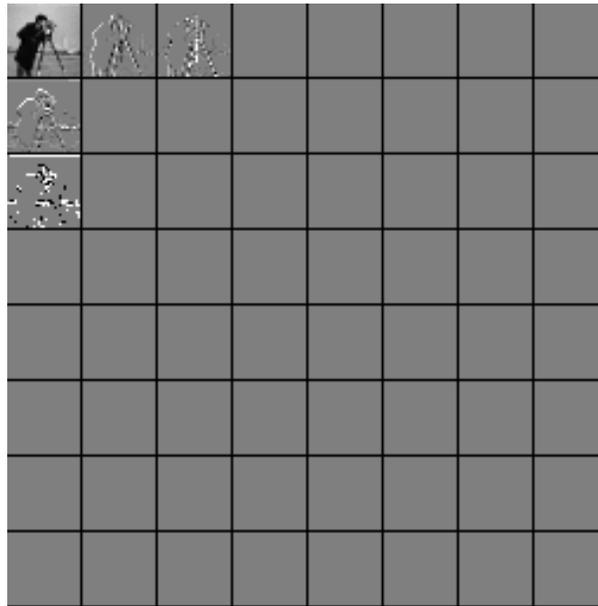
La norme Jpeg utilise en outre un codage entropique qui n'est pas pris en compte ci-dessous :

originale



256x256 pixels codés
sur 8 bits.

DCT2d quantifiée



DCT2d inverse



Bitrate : 0,25

Size : 16 Ko

Mean Square Error : 378,7

Transmission progressive d'image correspondant à une allocation de débit variable du codage DCT2d.

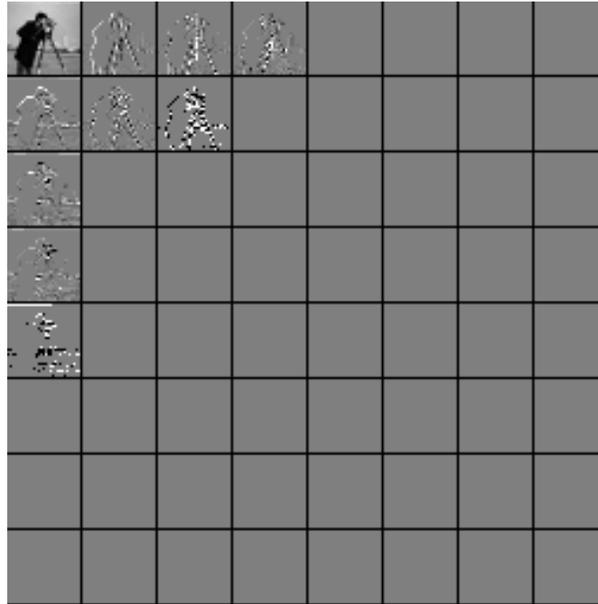
La norme Jpeg utilise en outre un codage entropique qui n'est pas pris en compte ci-dessous :

originale



256x256 pixels codés
sur 8 bits.

DCT2d quantifiée



DCT2d inverse



Bitrate : 0,5
Size : 32 Ko
Mean Square Error : 239,5

Transmission progressive d'image correspondant à une allocation de débit variable du codage DCT2d.

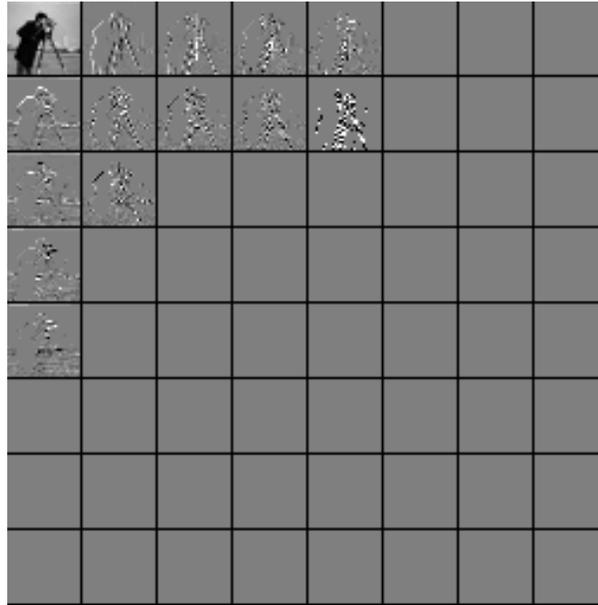
La norme Jpeg utilise en outre un codage entropique qui n'est pas pris en compte ci-dessous :

originale



256x256 pixels codés
sur 8 bits.

DCT2d quantifiée



DCT2d inverse



Bitrate : 0,75
Size : 48 Ko
Mean Square Error : 171,8

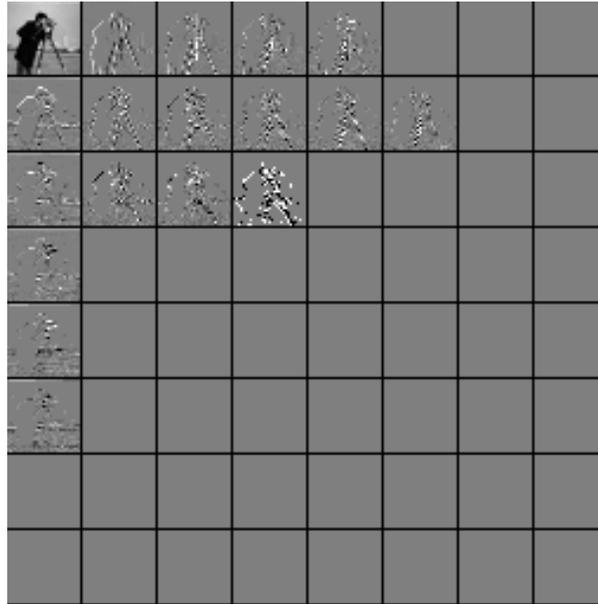
Transmission progressive d'image correspondant à une allocation de débit variable du codage DCT2d.

La norme Jpeg utilise en outre un codage entropique qui n'est pas pris en compte ci-dessous :

originale

DCT2d quantifiée

DCT2d inverse



256x256 pixels codés
sur 8 bits.

Bitrate : 1
Size : 64 Ko
Mean Square Error : 130,2

Transmission progressive d'image correspondant à une allocation de débit variable du codage DCT2d.

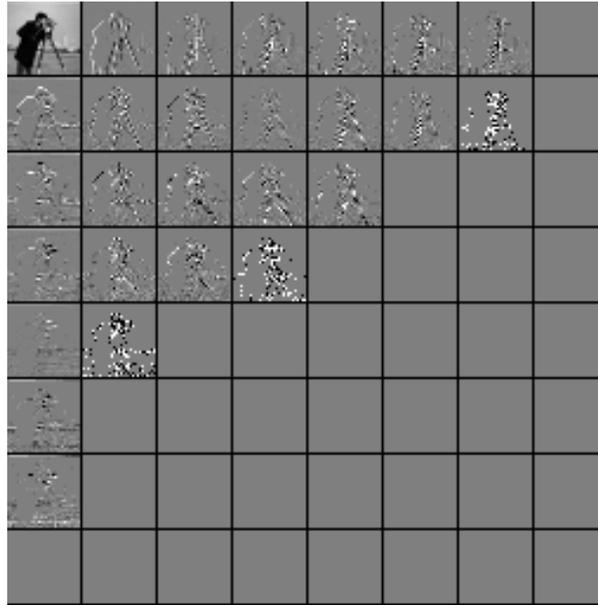
La norme Jpeg utilise en outre un codage entropique qui n'est pas pris en compte ci-dessous :

originale



256x256 pixels codés
sur 8 bits.

DCT2d quantifiée



DCT2d inverse



Bitrate : 1,5
Size : 96 Ko
Mean Square Error : 79,8

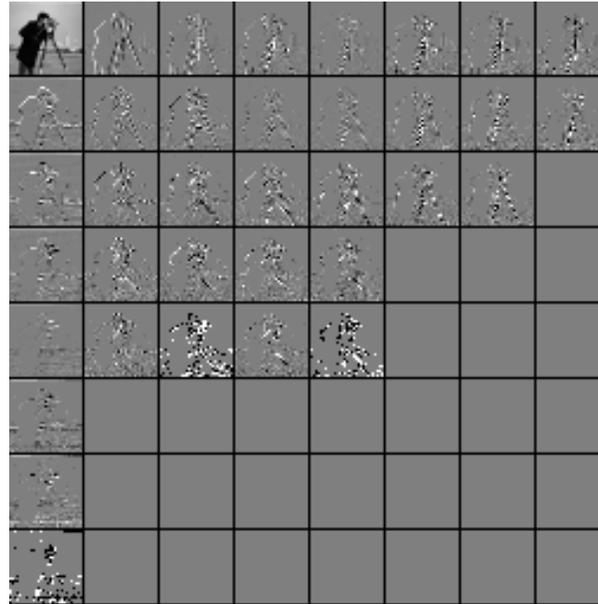
Transmission progressive d'image correspondant à une allocation de débit variable du codage DCT2d.

La norme Jpeg utilise en outre un codage entropique qui n'est pas pris en compte ci-dessous :

originale

DCT2d quantifiée

DCT2d inverse



256x256 pixels codés
sur 8 bits.

Bitrate : 2
Size : 128 Ko
Mean Square Error : 50,0

Transmission progressive d'image correspondant à une allocation de débit variable du codage DCT2d.

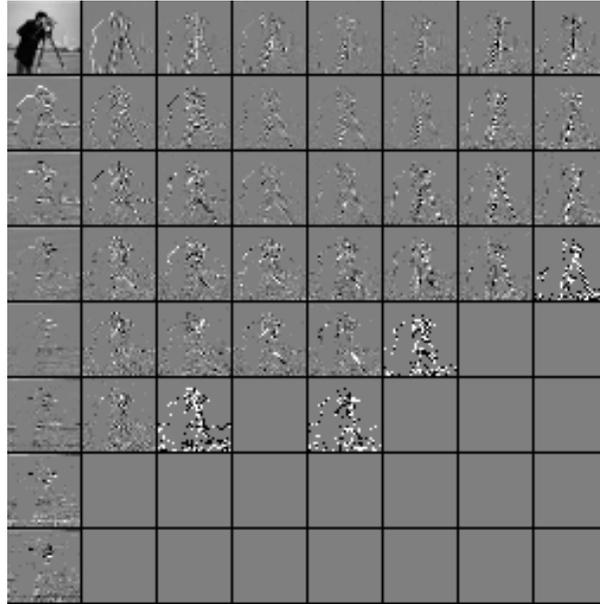
La norme Jpeg utilise en outre un codage entropique qui n'est pas pris en compte ci-dessous :

originale



256x256 pixels codés
sur 8 bits.

DCT2d quantifiée



DCT2d inverse



Bitrate : 2,5
Size : 160 Ko
Mean Square Error : 31,6

Transmission progressive d'image correspondant à une allocation de débit variable du codage DCT2d.

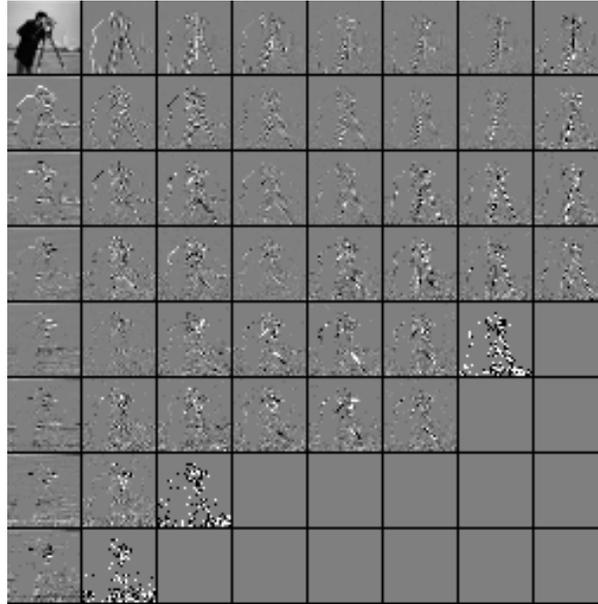
La norme Jpeg utilise en outre un codage entropique qui n'est pas pris en compte ci-dessous :

originale



256x256 pixels codés
sur 8 bits.

DCT2d quantifiée



DCT2d inverse



Bitrate : 3
Size : 192 Ko
Mean Square Error : 19,9

Transmission progressive d'image correspondant à une allocation de débit variable du codage DCT2d.

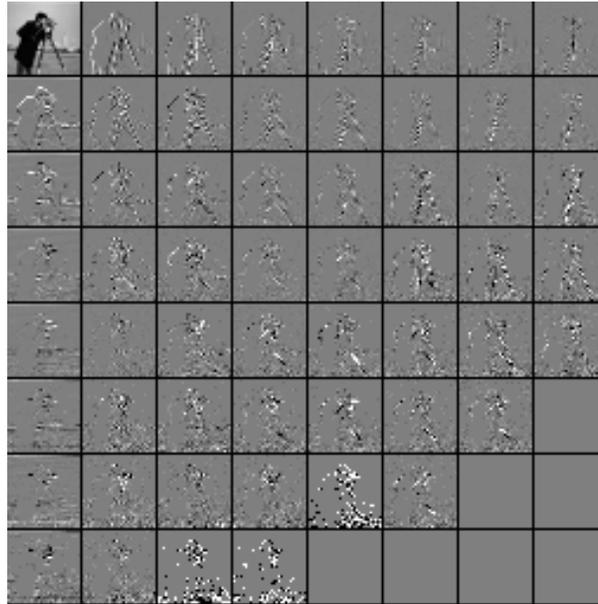
La norme Jpeg utilise en outre un codage entropique qui n'est pas pris en compte ci-dessous :

originale



256x256 pixels codés
sur 8 bits.

DCT2d quantifiée



DCT2d inverse



Bitrate : 3,5
Size : 224 Ko
Mean Square Error : 12,4

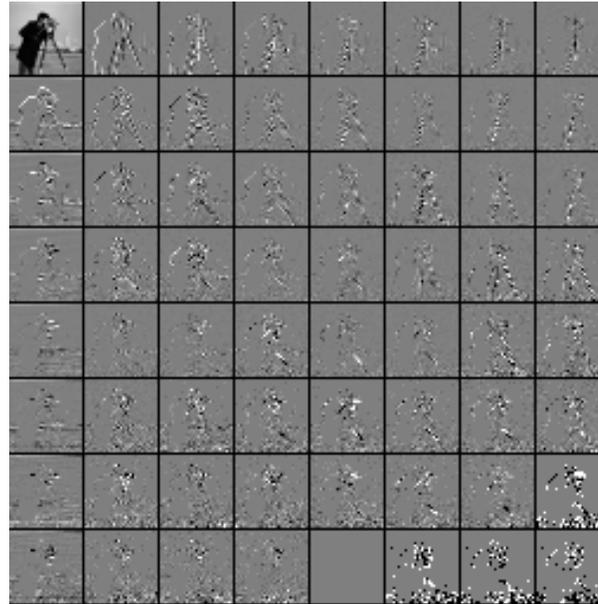
Transmission progressive d'image correspondant à une allocation de débit variable du codage DCT2d.

La norme Jpeg utilise en outre un codage entropique qui n'est pas pris en compte ci-dessous :

originale

DCT2d quantifiée

DCT2d inverse

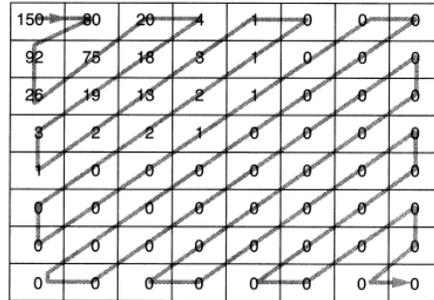


256x256 pixels codés
sur 8 bits.

Bitrate : 4
Size : 256 Ko
Mean Square Error : 7,5

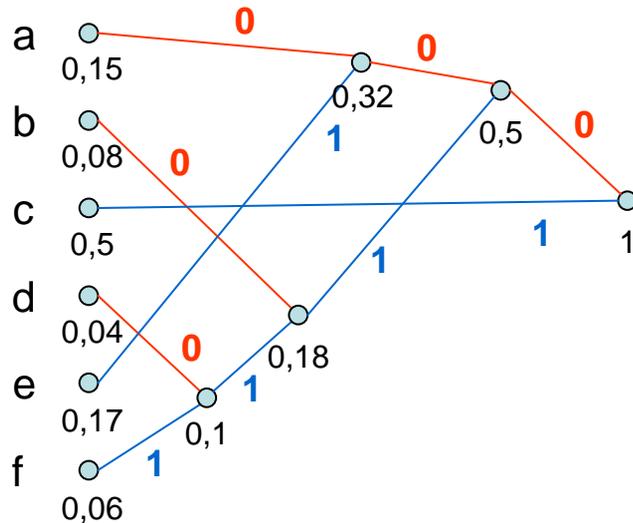
Après la transformée en DCT2d et la quantification, la deuxième partie de la norme Jpeg utilise une *compression sans perte* qui va : (1) regrouper les nombreuses valeurs nulles par plages (*Run Length Encoding*), et (2) exploiter le déséquilibre dans la distribution des valeurs à coder en attribuant à chaque valeur (symbole) une longueur (nombre de bits) inversement proportionnelle à sa probabilité d'apparition (*codage entropique*).

(1) Chaque bloc DCT2d quantifié est transformé en un vecteur ligne grâce à un parcours en zig-zag :



Cette transformation regroupe les valeurs faibles en fin de vecteur.

(2) Les valeurs non nulles sont codées par le codage de Huffman :



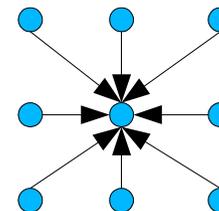
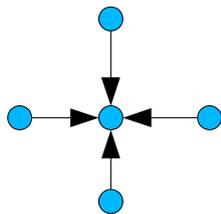
Symbole	Probab.	Code
a	0,15	000
b	0,08	010
c	0,5	1
d	0,04	0110
e	0,17	001
f	0,06	0111

III-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.

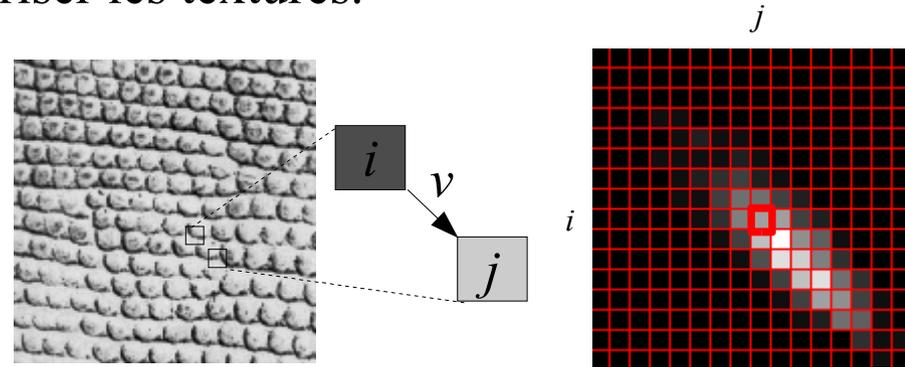
Par exemple, le modèle des champs de Markov considère l'image comme la réalisation d'un champ aléatoire (chaque pixel correspondant à une variable aléatoire), où la valeur prise par un pixel ne dépend que de celle de ses voisins (selon une topologie discrète donnée, voir plus loin).



III-3 : Le modèle statistique

Un autre exemple remarquable d'analyse statistique est l'utilisation des matrices de cooccurrence, 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 $(x, x+v)$.

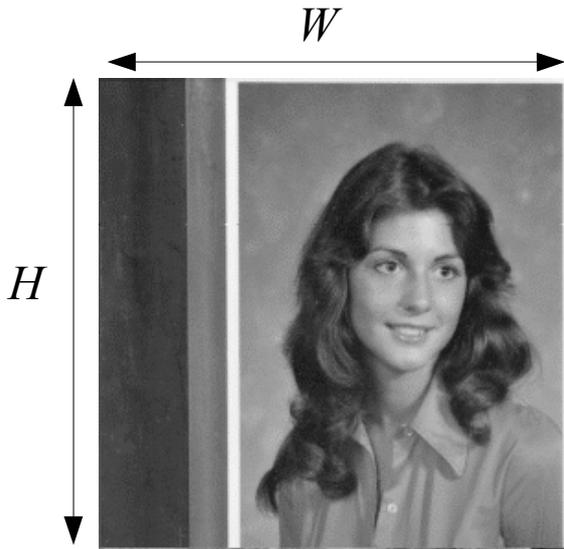


Ces techniques, et d'autres exemples de modèles statistiques seront présentées plus en détail dans les cours TERI et TDI. Dans le cadre de cette introduction, nous développerons seulement le premier outil statistique d'analyse des images : l'histogramme.

L'histogramme :

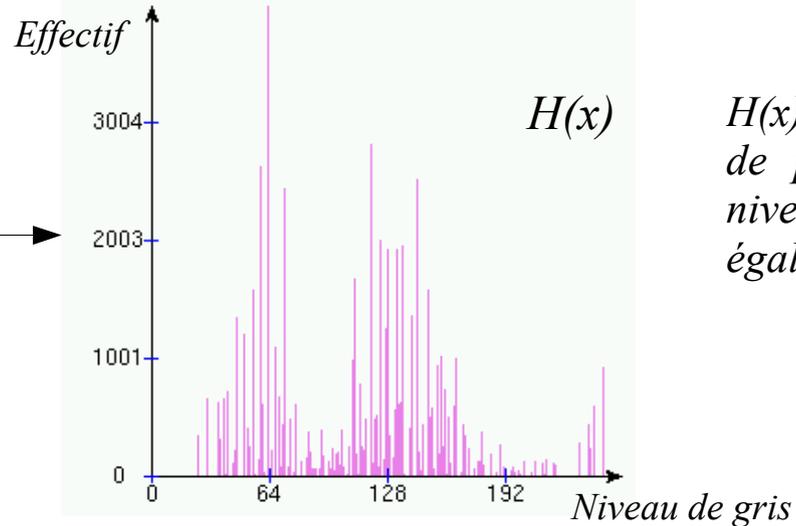
Outil de base pour l'étude des capteurs ou de la dynamique d'une scène, il est utilisé par certains opérateurs d'analyse. On retiendra cependant qu'il ne faut pas considérer l'histogramme comme une caractéristique fondamentale de l'image dans la mesure où on peut le transformer radicalement sans changer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

Histogramme



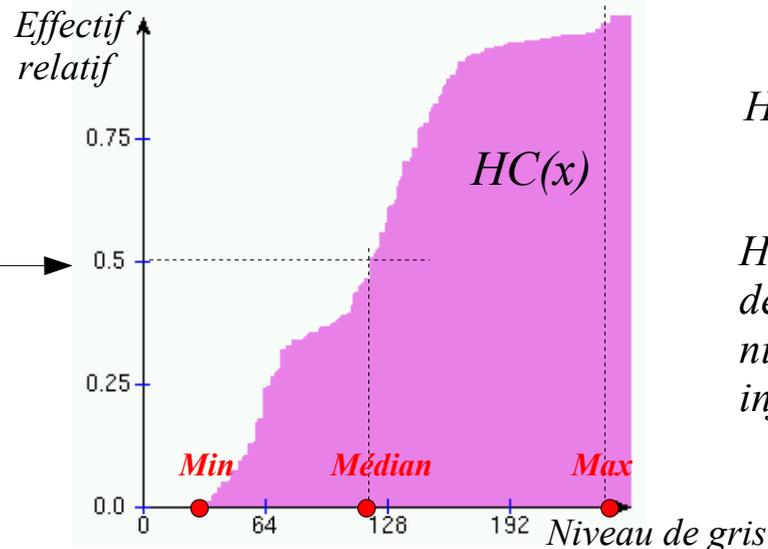
L'histogramme représente la répartition des pixels en fonction de leur niveau de gris. Il fournit diverses informations comme les statistiques d'ordre (voir ci-contre), l'entropie (voir précédemment), et peut permettre d'isoler des objets.

Histogramme



$H(x)$ est le nombre de pixels dont le niveau de gris est égal à x .

Histogramme cumulatif normalisé



$$HC(x) = \frac{\sum_{i=0}^x H(x)}{W \times H}$$

$HC(x)$ est le taux de pixels dont le niveau de gris est inférieur à x .

Traitement à base d'histogramme

On présente dans la suite quelques traitements 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.

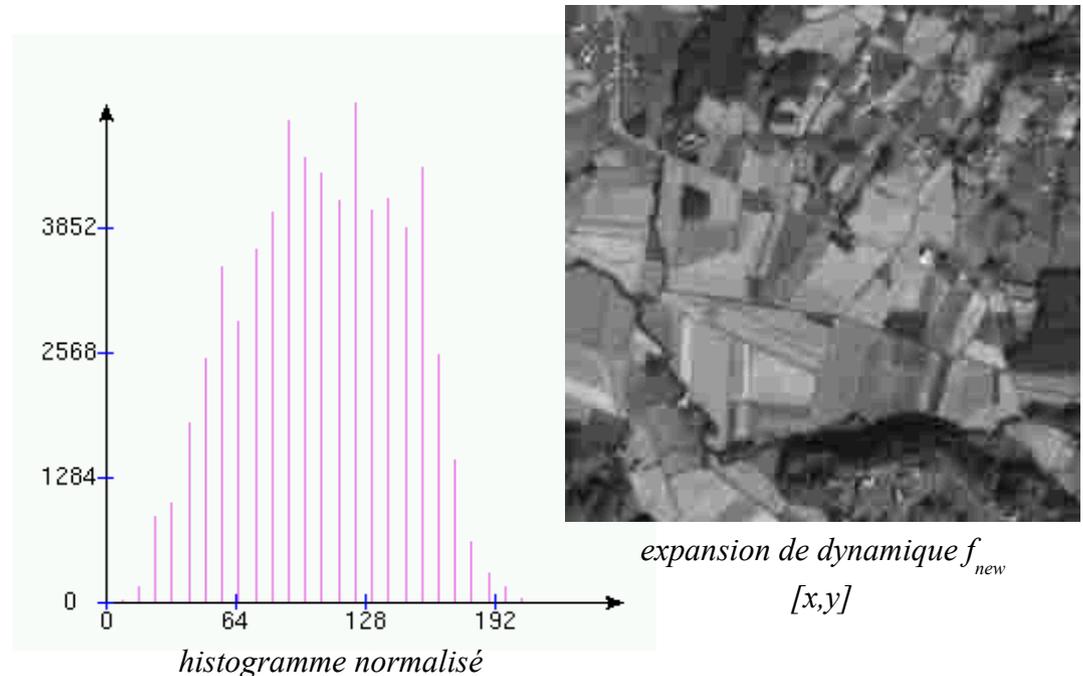
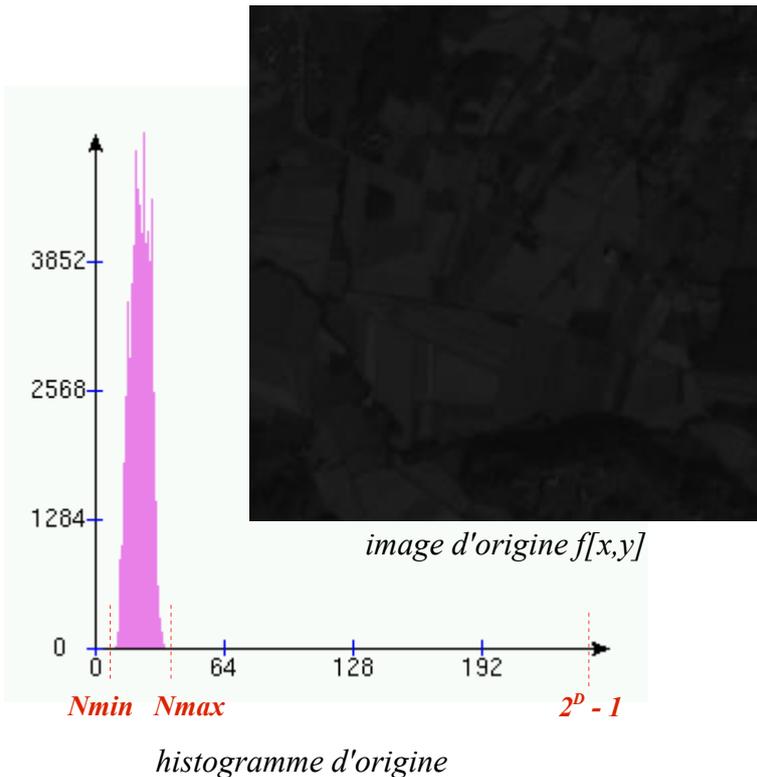
- D : dynamique
- $Nmin$: la plus petite valeur dans l'image
- $Nmax$: la plus grande valeur dans l'image

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

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

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

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



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*).

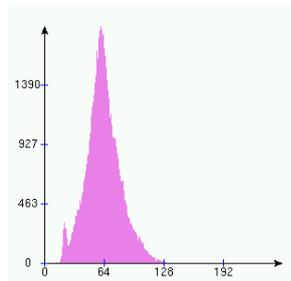
• La technique classique consiste à rendre « le plus linéaire possible » l'histogramme cumulé de l'image en utilisant la transformation suivante :

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

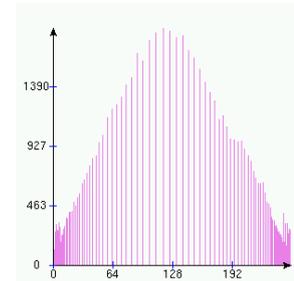
- D : dynamique
- (w, h) : dimension de l'image
- $HC(.)$: histogramme cumulé



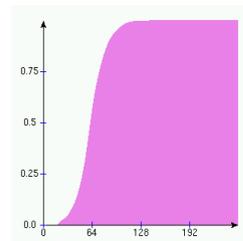
Original $f[x, y]$



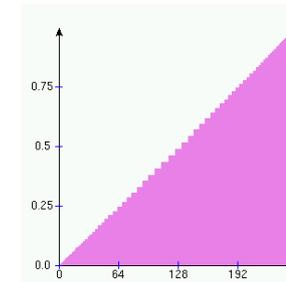
Histogramme de f



Histogramme de f_{new}



Histogramme cumulé de f



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'imager infrarouge.

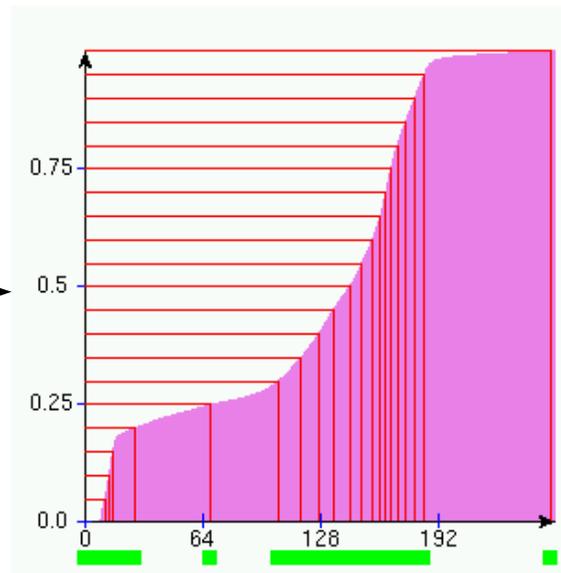
Histogramme : segmentation

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*

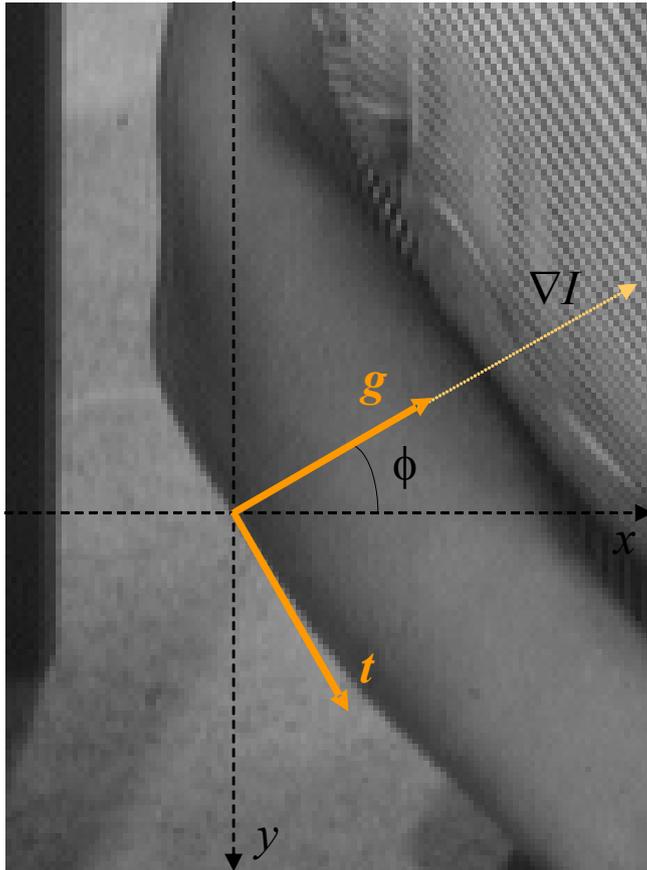


Image segmentée

III-4 : 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.

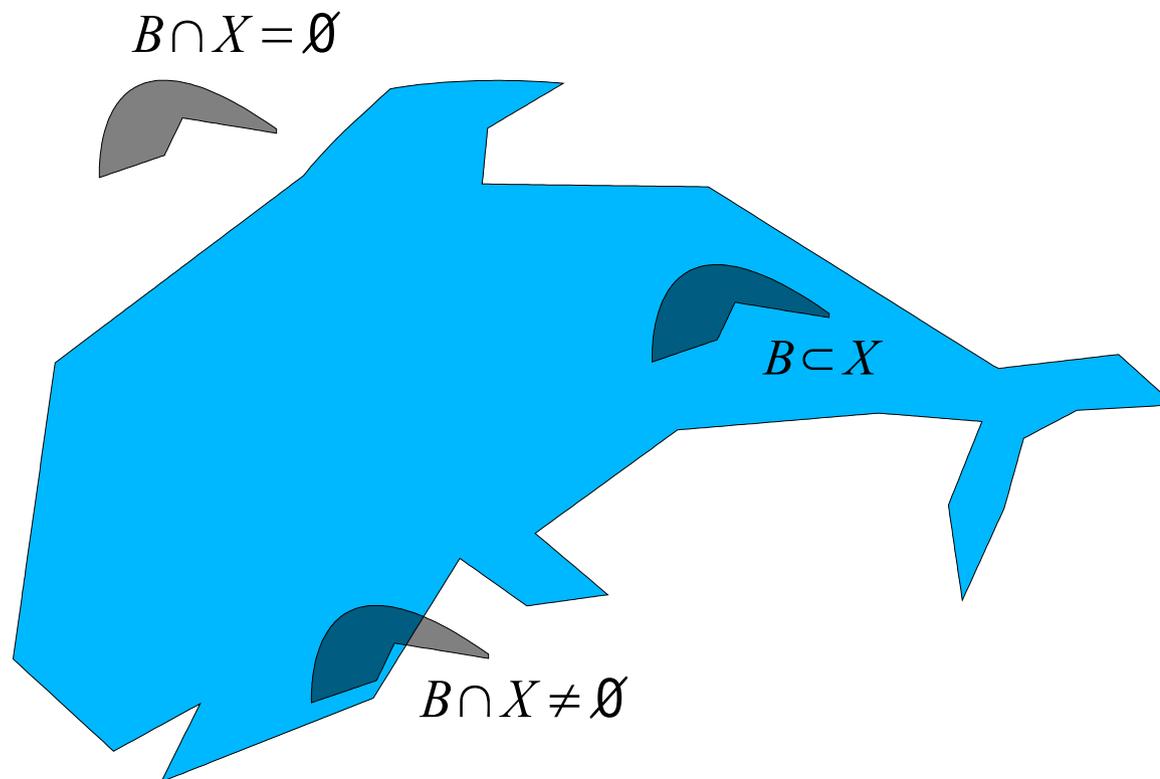


Au premier ordre, on peut ainsi associer à chaque point (x,y) un repère propre (t,g) , où le vecteur t donne la direction de l'*isophote* (ligne de variation minimale) et g la direction orthogonale, celle du *gradient*.

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 résolutions (voirs prochains cours TERI et TDI).

III-5 : 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).



Erosion 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)



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

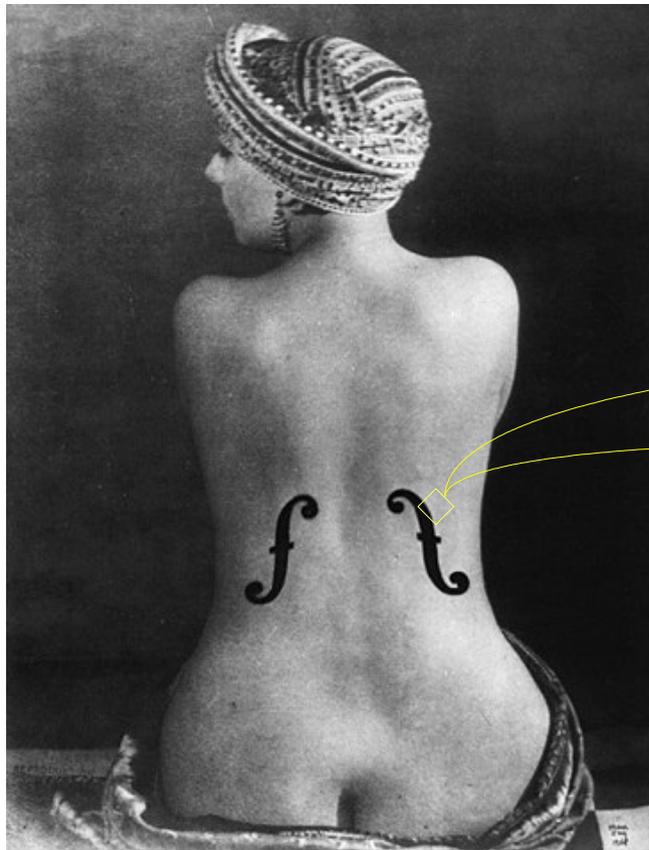


$$\delta_B(X) = \{x \in \mathbb{R}^2 ; B_x \cap X \neq \emptyset\}$$

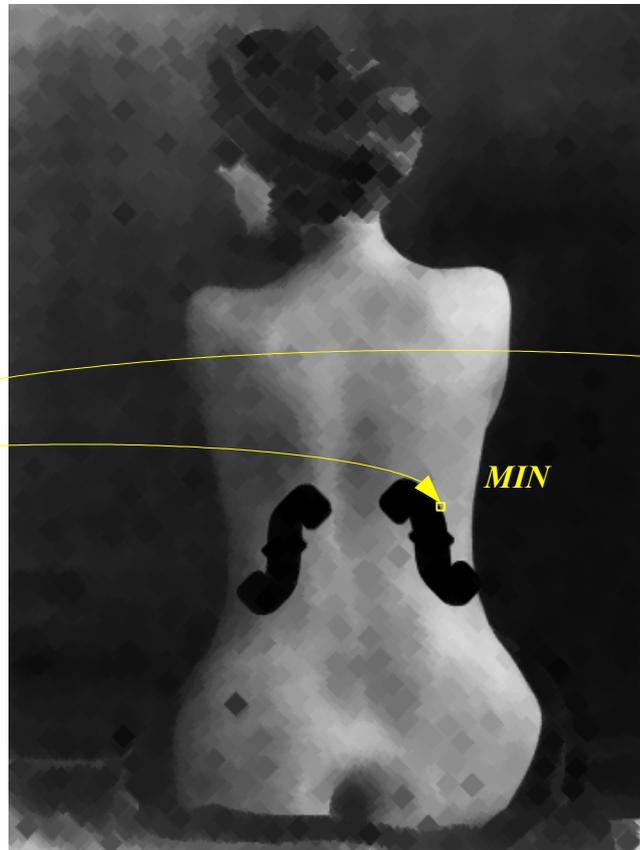
(élément structurant : disque)

Erosion et Dilatation

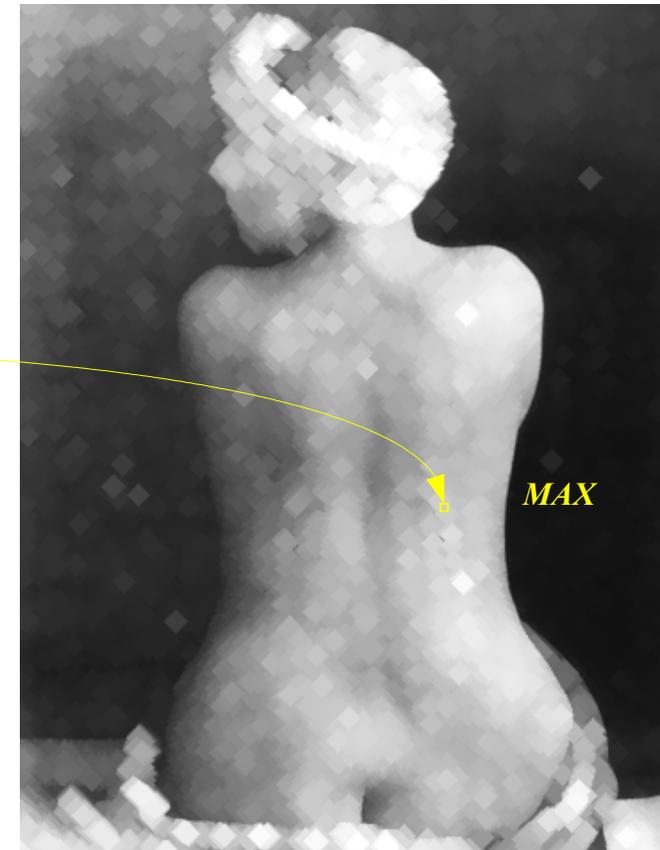
L'érosion et la dilatation, et par suite l'ensemble des transformations morphologiques, se généralisent des ensembles (Images binaires) aux fonctions (Images en niveaux de gris) par l'intermédiaire des ensembles de niveaux : $I_n = \{x \in \mathbb{R}^2 ; I(x) \leq n\}$



Original (Man Ray - 1924)



$\varepsilon_B(I)$

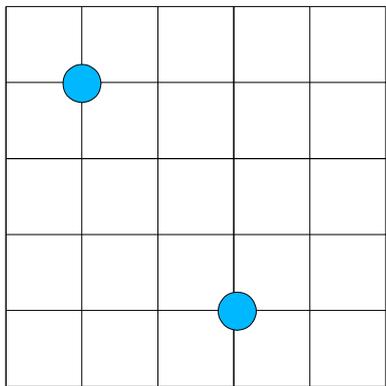


$\delta_B(I)$

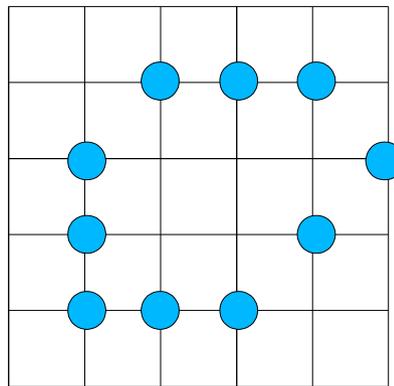
(élément structurant : losange)

III-6 : Le modèle discret

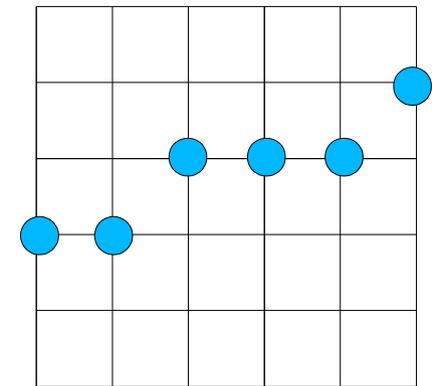
La *géométrie discrète* est une discipline au moins aussi ancienne que le traitement d'images. Alors que le modèle différentiel considère les structures géométriques (courbes, surfaces, droites, ...) comme des approximations numériques de leurs homologues continues, ou que le modèle fréquentiel traduit la discrétisation 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 ?



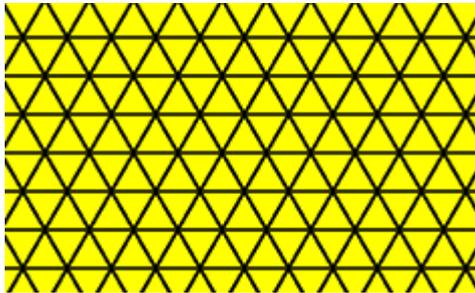
Qu'est-ce qu'un trou ?



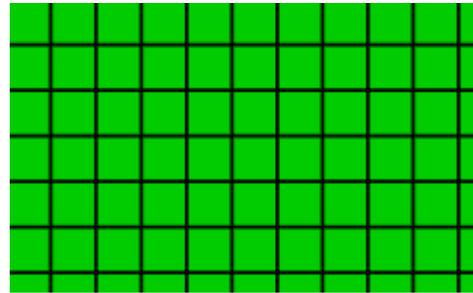
Qu'est-ce qu'une droite ?

Pavages du plan

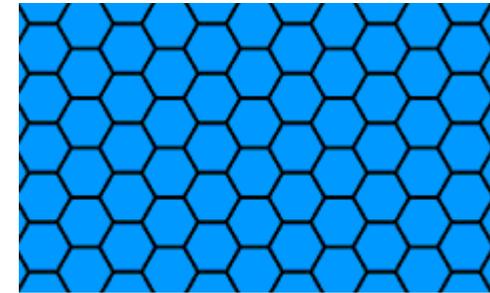
Un pavage du plan est une partition du plan en cellules élémentaires (pixels).
Il n'existe que 3 pavages réguliers du plan :



triangulaire

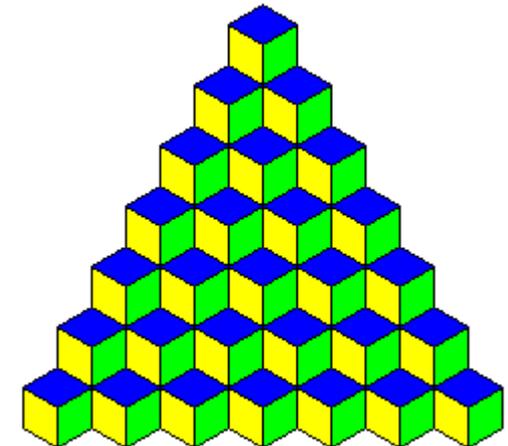
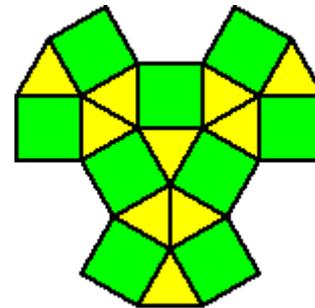
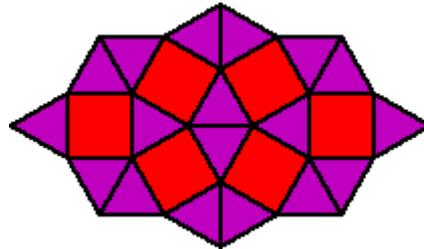
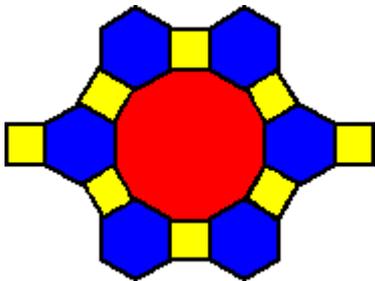


carré



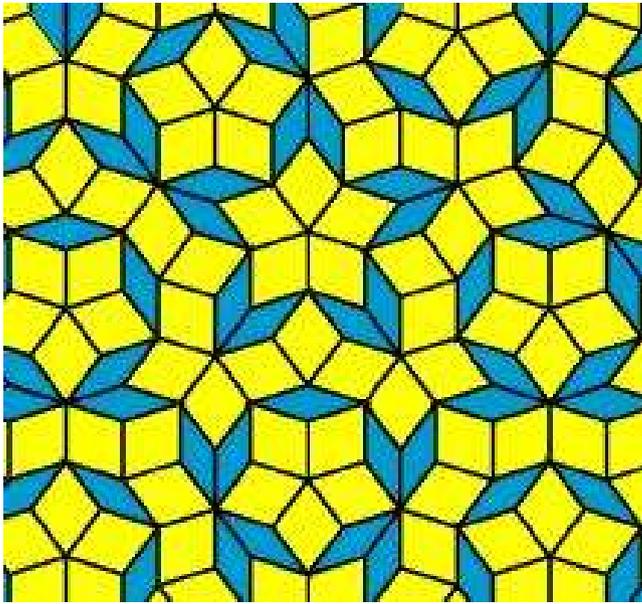
hexagonal

... mais de nombreux pavages irréguliers :

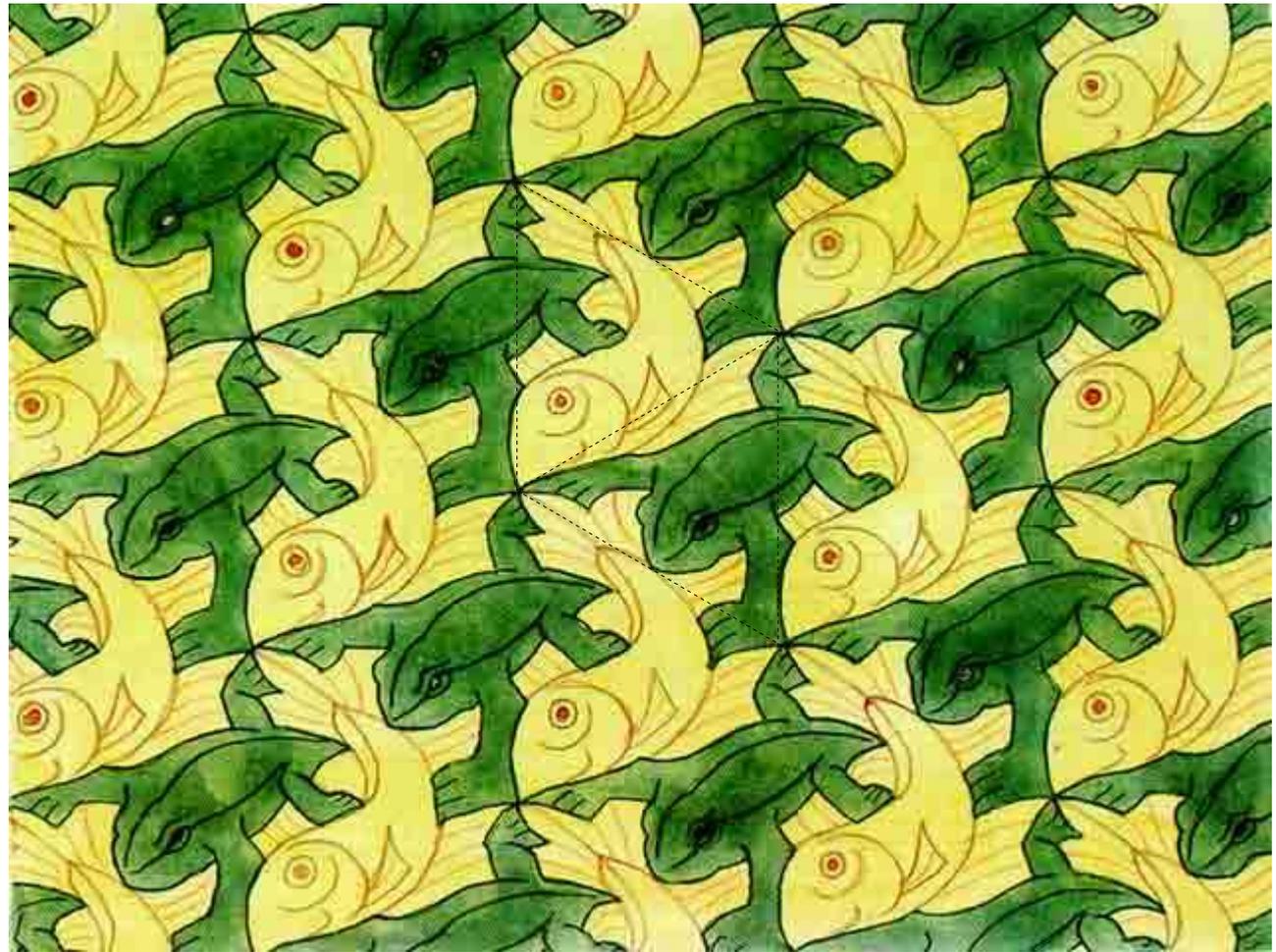


Pavages du plan

D'autres pavages irréguliers du plan...



Pavage aperiodique de Penrose

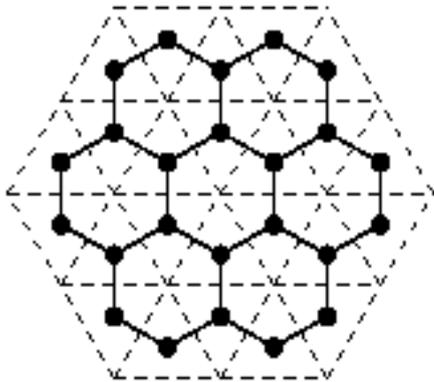


Pavage périodique d'Escher

Pavages et maillages

A tout pavage du plan on peut associer un *graphe* où les *sommets* (noeuds) 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.

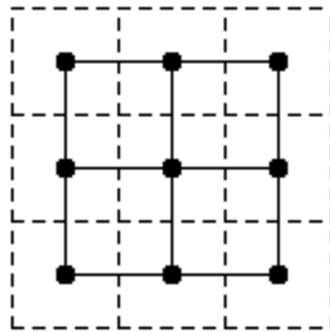
Les pavages et les maillages réguliers sont *duaux* :



Pavage triangulaire



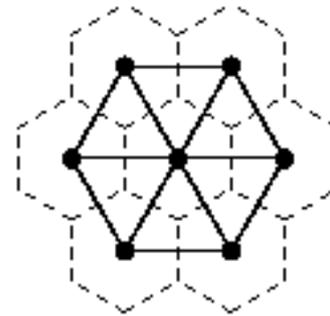
Maillage hexagonal



Pavage carré



Maillage carré



Pavage hexagonal



Maillage triangulaire

Questions :

- représentation dans Z^2 ?
- combien de directions ?
- récursivité ?

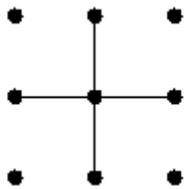
Maillage et connexité

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.

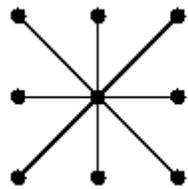
$$X \subset \mathbb{Z}^2 ; S \subset X^2$$

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

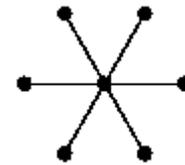
$$x \approx y \Leftrightarrow (x, y) \in S$$



maille carrée
4-connexe



maille carrée
8-connexe



maille triang.
6-connexe

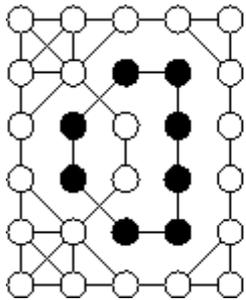
La clôture transitive de la relation de voisinage est une relation d'équivalence
« il existe un chemin connexe entre x et y » :

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

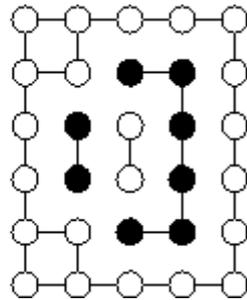
Les classes d'équivalence de cette relation s'appellent les *composantes connexes* de X

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...



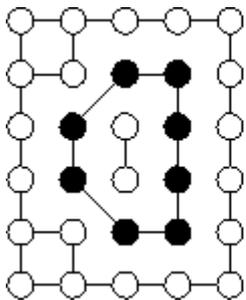
8-connectivité



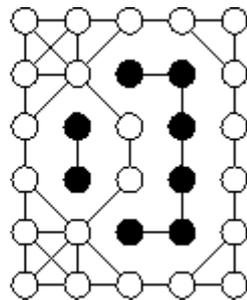
4-connectivité

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 :



(8,4)-connectivité



(4,8)-connectivité

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



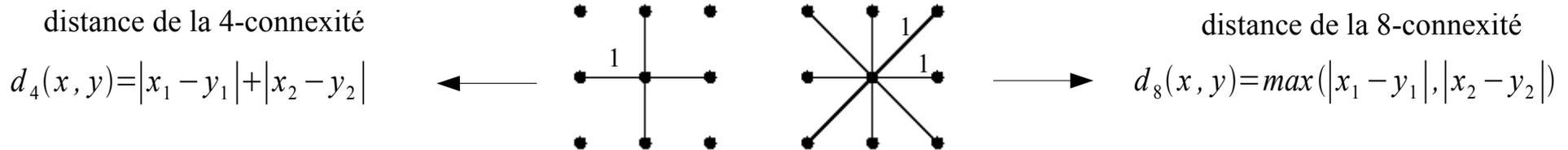
Questions :

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

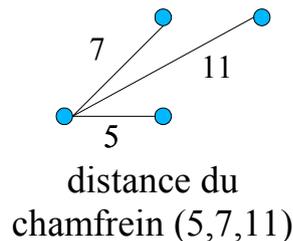
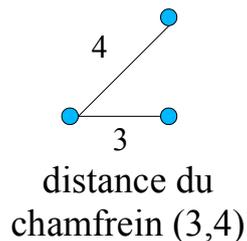
- en (8,4)-connectivité ?
- en (4,8)-connectivité ?

Métrie 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 valant toutes les arêtes du maillage par la valeur 1, on trouve :

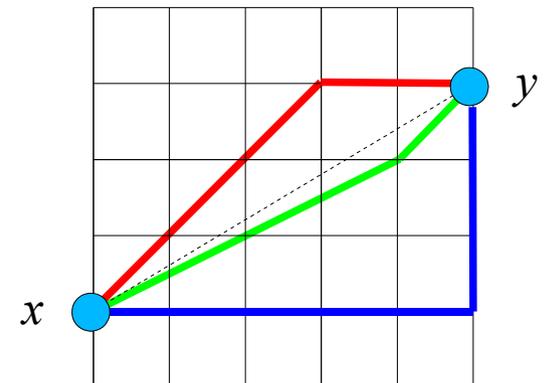


On peut aussi valuer différemment les arêtes du maillage 8-connexé, voire utiliser des maillages plus complexes (i.e. des voisinages plus grands) :



Questions :

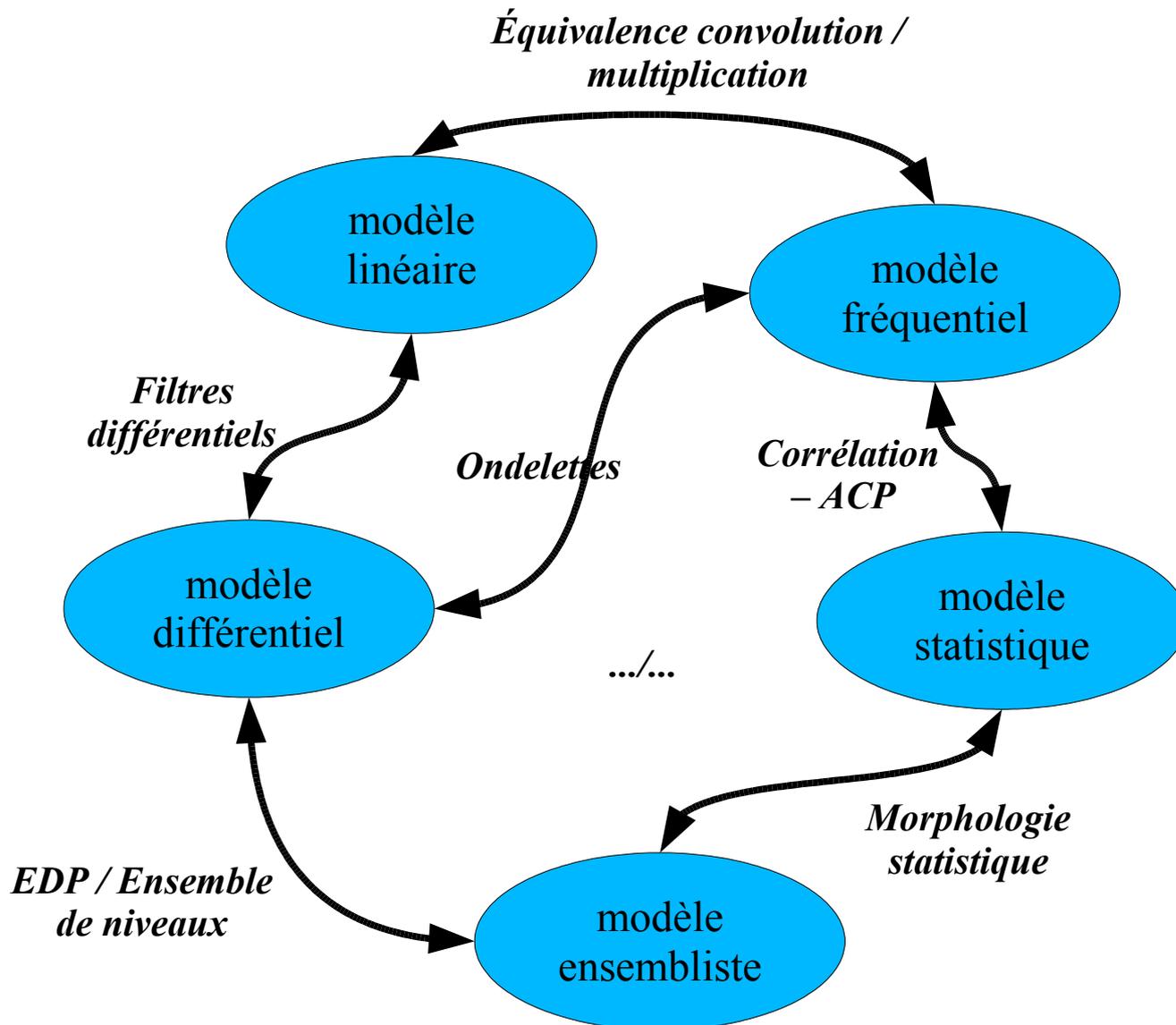
calculer les distances $d_4(x, y)$, $d_8(x, y)$, $d_{ch(3,4)}(x, y)$, $d_{ch(5,7,11)}(x, y)$ entre les 2 points x et y ci-contre :



Conclusion

A retenir pour ce cours :

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



Sources et bibliographie - Liens utiles

Livres :

- J.P. Cocquerez et S. Philipp « *Analyse d'images : filtrage et segmentation* » Masson 1995
- R.C. Gonzalez et Woods « *Digital Image Processing - 2d edition* » Addison Wesley 2002
- A. Rosenfeld et A.C. Kak « *Digital picture processing* » Academic Press - London 1982.
- H. Maître (ss la direction de) « *Le traitement des images* » Hermes Lavoisier IC2 2003.
- J.R. Parker « *Algorithms for Image Processing and Computer Vision* » Wiley & Sons 1997.
- S. Bres, J.M. Jolion, F. Lebourgeois « *Traitement et analyse des images numériques* » Hermes Lavoisier 2003
- I.T. Young, J.J. Gerbrands et L.J. Van Vliet « *Fundamentals of Image Processing* » Université de Delft. (sur internet : <http://www.ph.tn.tudelft.nl/~lucas/publications/1995/FIP95TYJGLV/FIP2.2.pdf>)
- D. Lingrand « *Introduction au Traitement d'images* » Vuibert 2004

Pages web :

- Telesun – INSA : <http://telesun.insa-lyon.fr/~telesun/>
- Univ. de Delft : <http://www-ict.its.tudelft.nl/html/education/courses/>
- Projet Marble : <http://www.icbl.hw.ac.uk/marble/vision/low/fundamentals/intro.htm>
- CVOnline - Univ. d'Édimbourg : <http://www.dai.ed.ac.uk/CVonline/>