

Traitement d'images et vision artificielle

Cours
ENSTA
D9-2

Antoine Manzanera - Unité d'Électronique et d'Informatique



TI & Vision - Introduction

Ce cours a pour objectif la découverte de techniques parmi les plus récentes en traitement d'images et en vision par ordinateur.

Par *traitement d'images*, on désigne l'ensemble des *opérations sur les images* numériques, qui transforment une image en une autre image, ou en une autre primitive formelle.

La *vision par ordinateur* désigne la *compréhension* d'une scène ou d'un phénomène à partir d'informations « image », liant intimement *perception*, *comportement* et *contrôle*.

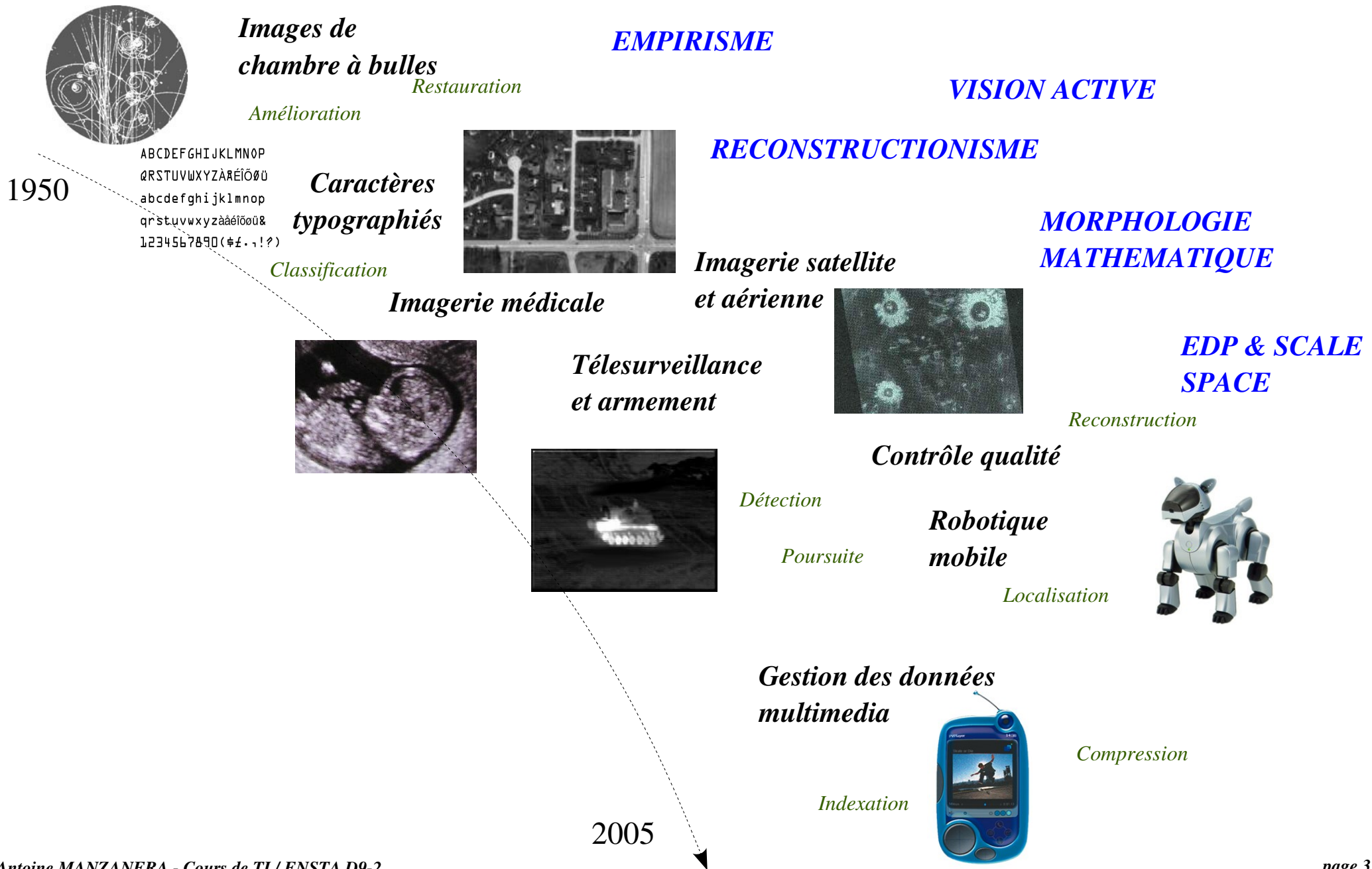
Les domaines traités vont du *traitement du signal* à l'*intelligence artificielle*, on ne saurait donc prétendre à l'exhaustivité, mais on vise plutôt l'exploration d'un certain nombre de techniques importantes et actuelles.

TI et vision sont des disciplines relativement jeunes (~années 60) et qui évoluent rapidement. Elles sont en plein expansion et donnent lieu chaque année à une profusion de travaux, académiques, technologiques, industriels.

Cette profusion s'explique par le caractère ardu du sujet : *complexité algorithmique* dû aux énormes volumes de données, caractère *mal posé* des problèmes et difficultés à formaliser une *faculté biologique* « évidente ».

D'autre part l'engouement pour ces disciplines s'explique par la multiplication permanente d'*applications* et d'*enjeux industriels* dans des domaines aussi variés que : médecine, télécommunications, automobile, météorologie, défense, jeux video, art, écologie...

Historique du Traitement d'Images



Applications du TI - 1

ROBOTIQUE

PRAXITÈLE (INRIA)



conduite automatique

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

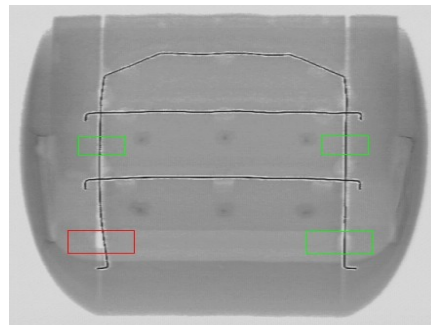
AIBO (SONY)



suivi et préhension
d'objets en robotique
mobile

CONTRÔLE INDUSTRIEL

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

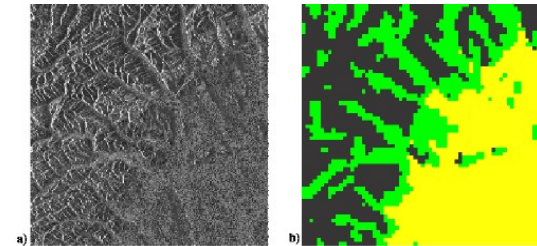


détection de défaut

- * environnement maîtrisé

TELEDETECTION

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



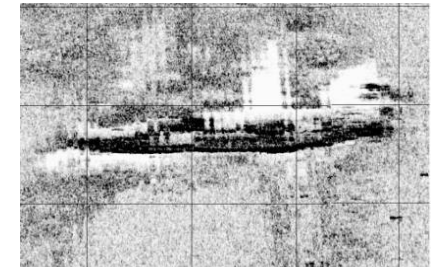
environnement

*Image d'un carrefour
(Univ. Karlsruhe)*



surveillance du
trafic routier

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



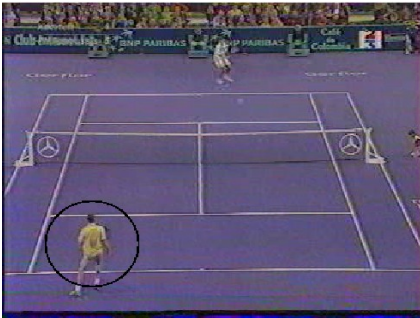
exploration des fonds marins

- * gros volume de données
- * temps différé possible

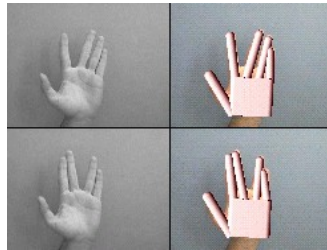
Applications du TI - 2

MULTIMEDIA

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



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



codage et compression

peinture du XVe siècle ternie et écaillée restaurée par filtrage médian et ajustement des couleurs (Univ. Québec)

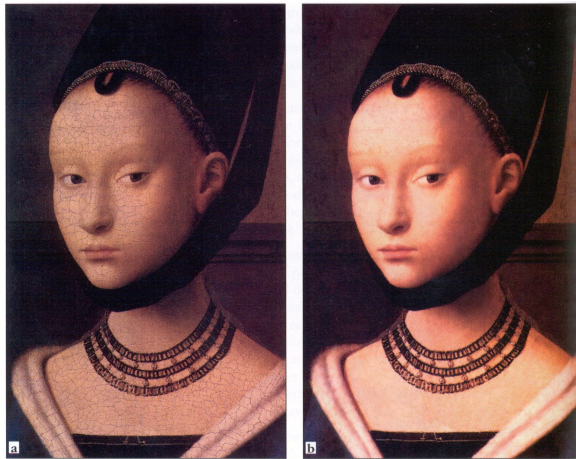


Figure 42. Portrait of a Lady, painted in 1470 by Petrus Christus:

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.

restauration d'images

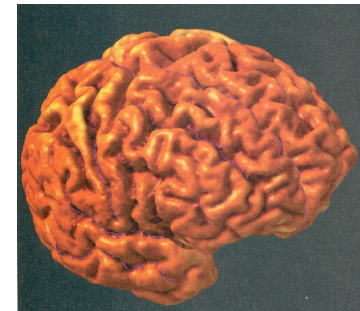
recherche d'images dans une base de données à partir d'une requête visuelle (ENSEA Cergy-Pontoise)



indexation d'images

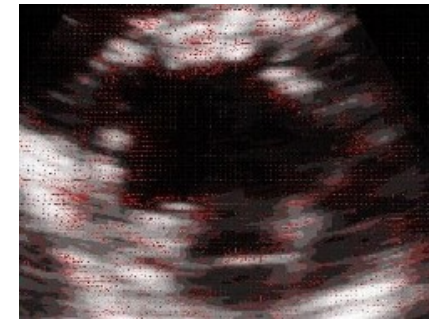
MÉDECINE

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



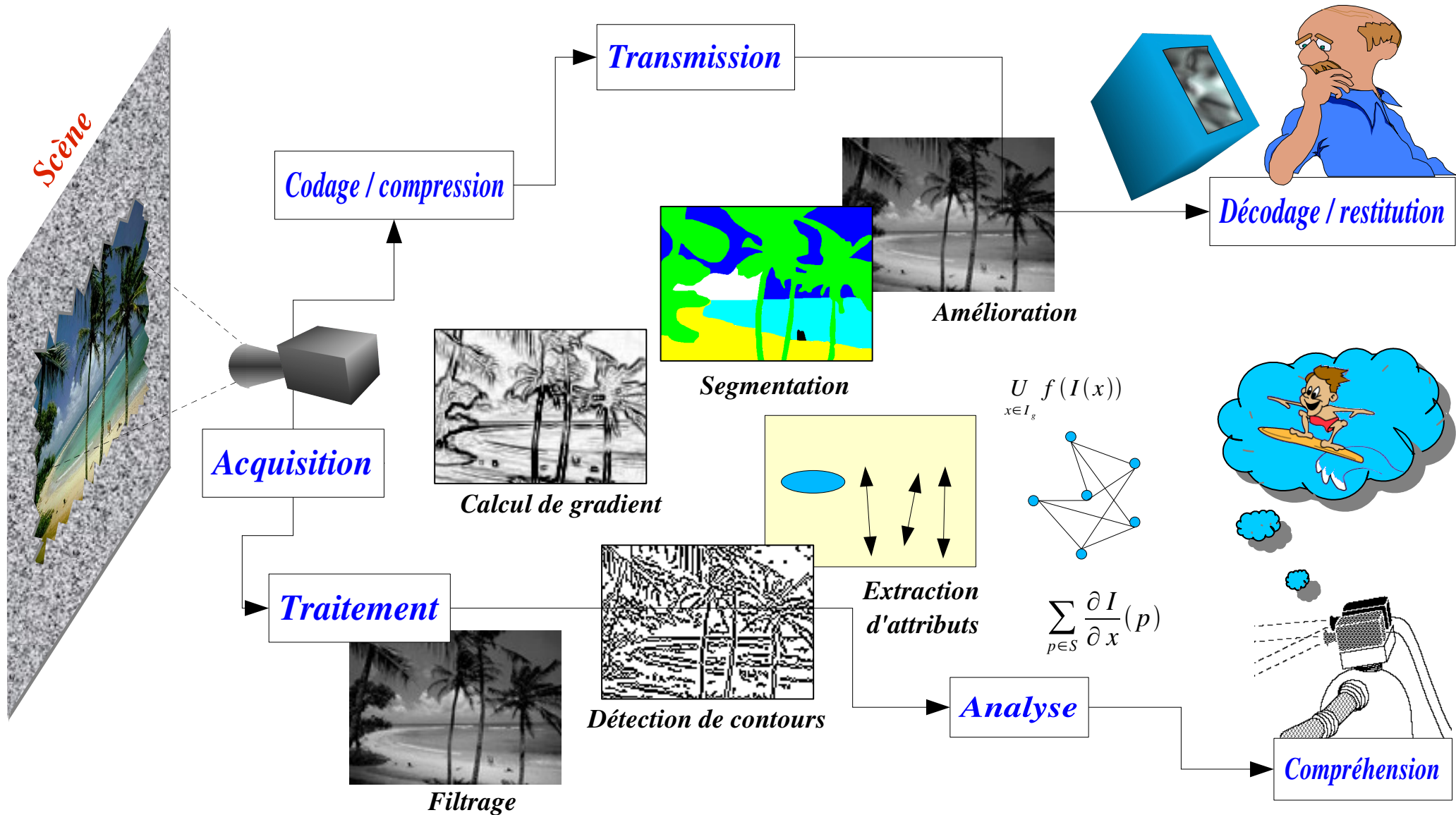
téléchirurgie et ChAO

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



Aide au diagnostic

Systemes à base de TI



Plan du cours D9-2

16 09 2010 Introduction / Filtrage (AM)

23 09 2010 Segmentation (JCB)

30 09 2010 TP1 : Filtrage, Contours (C)

07 10 2010 Espaces d'échelles / Couleur (AM)

21 10 2010 EDPs / Level sets + TP2 (PD)

04 11 2010 TP3 : Segmentation, Extraction de caractéristiques (C / Inti)

25 11 2010 Apprentissage et traitement d'images (BS)

09 12 2010 Mouvement / Estimation et Poursuite (AM)

16 12 2010 Vision 3d, Shape from shading, Stéréovision (RB+DF)

06 01 2011 Indexation et Catégorisation d'images et de vidéos (MC)

13 01 2011 TP4 : Catégorisation visuelle d'objets (Matlab)

20 01 2011 Fin des Tps + Vision active (JCB)

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>)

Pages web :

- CVOnLine - Univ. d'Édimbourg : <http://homepages.inf.ed.ac.uk/rbf/CVonline/>

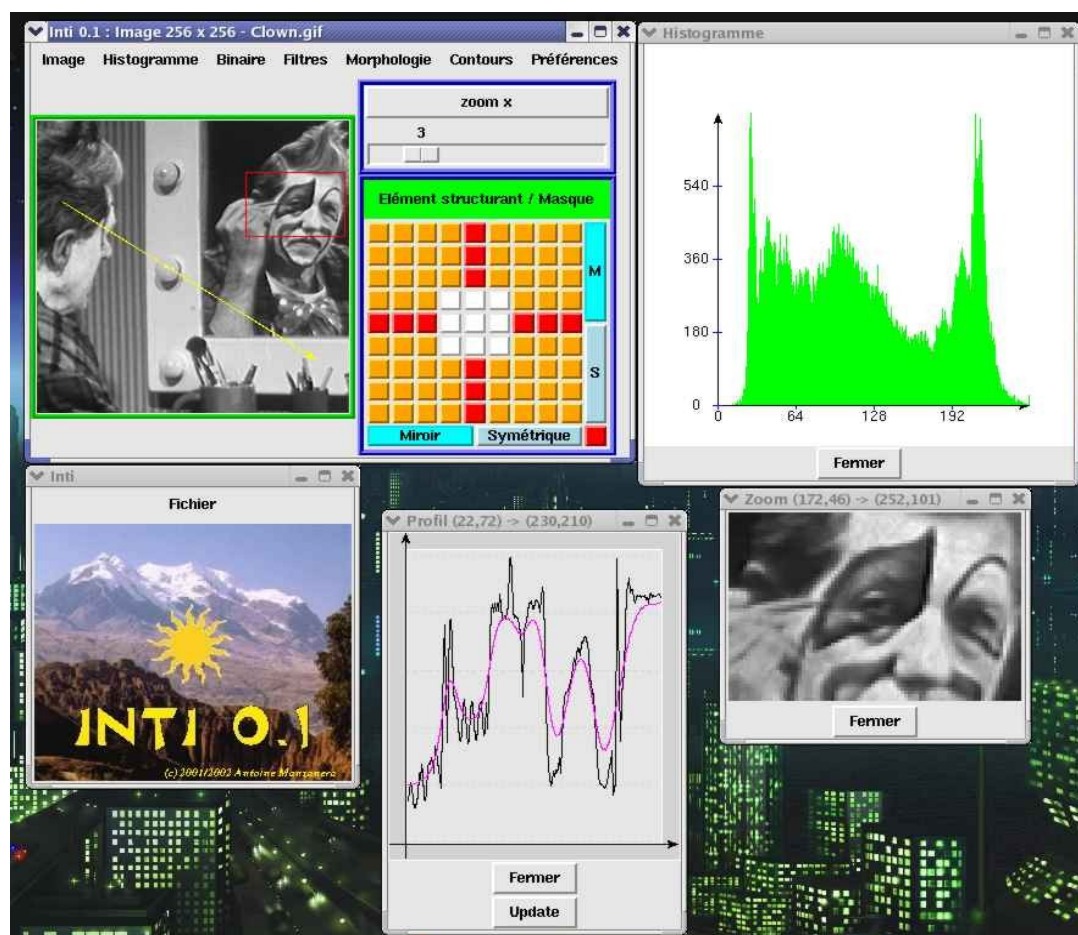
Logiciel support de cours

Beaucoup de traitements d'images présentés dans ce cours peuvent être expérimentés en utilisant le logiciel *Inti 0.1*, disponible sous trois versions différentes : (1) *Généraliste* (2) *Tutoriel de morphologie* (3) *Tutoriel Fourier 2d*.

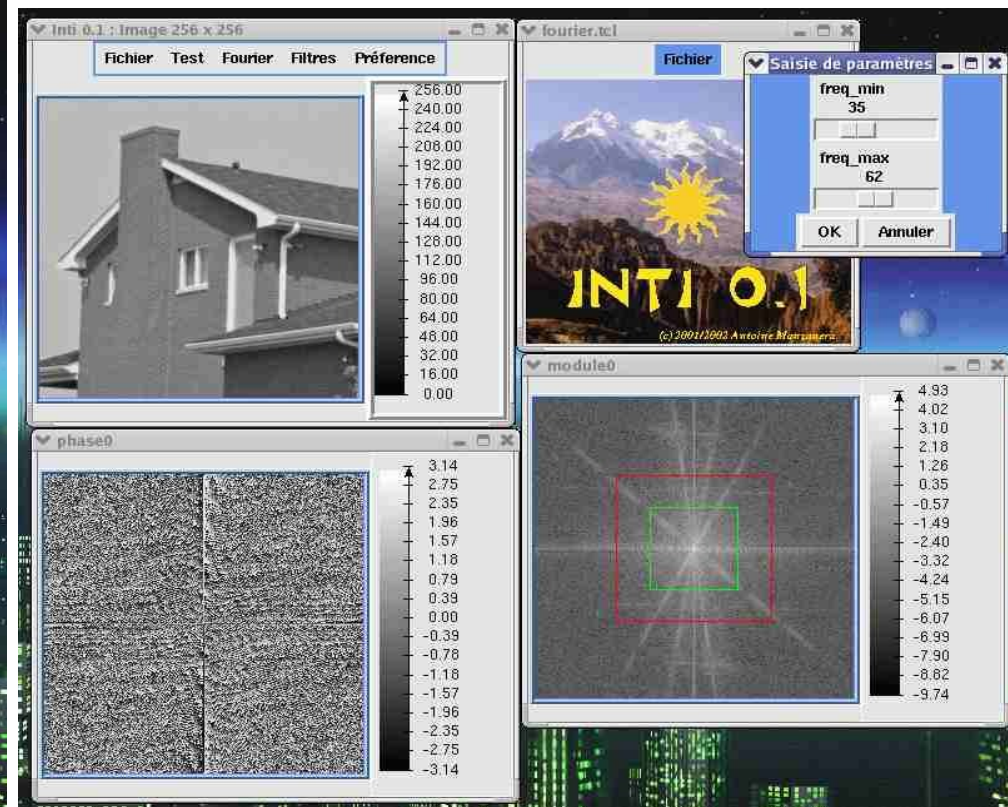


http://www.ensta.fr/~manzaner/Support_Cours.html

version généraliste :



version tutoriel Fourier 2d :



D9-2 : Plan du cours n°1

I] Introduction aux images numériques

- *Types d'images*
- *Vocabulaire*
- *Echantillonnage et quantification*
- *Outils fondamentaux*

II] Traitements à base d'histogramme

III] Filtres de lissage

- *Filtrage dans le domaine de Fourier*
- *Filtrage par convolution*
- *Implantation des filtres linéaires*
- *Filtres non linéaires*

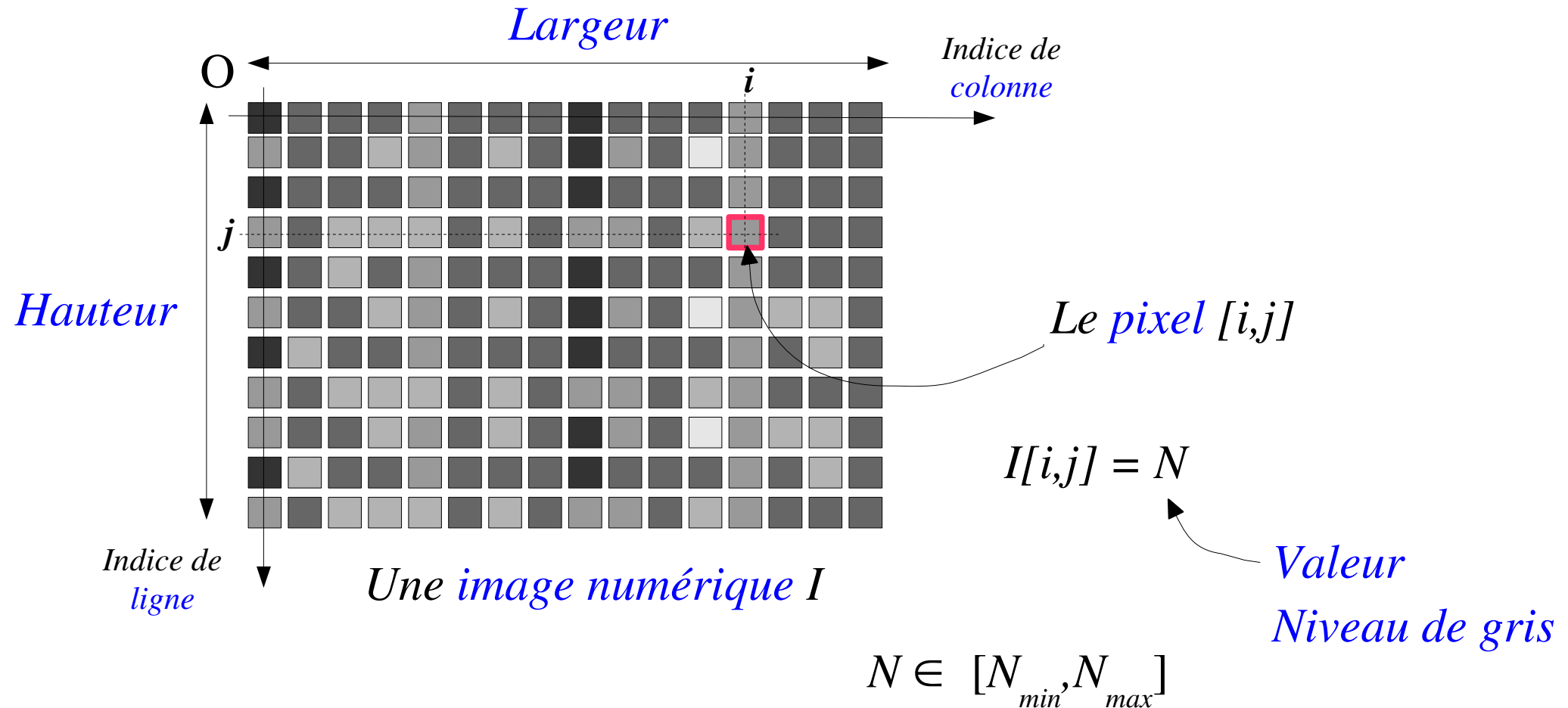
IV] Filtres dérivateurs

- *Filtrage dans le domaine de Fourier*
- *Filtrage par convolution*

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,...

I-2 Vocabulaire (1)



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

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

I-2 Vocabulaire (2)

Résolution...

...spatiale :

Échantillonnage



256x256



128x128



64x64



32x32

...tonale :

Quantification



6 bits



4 bits



3 bits



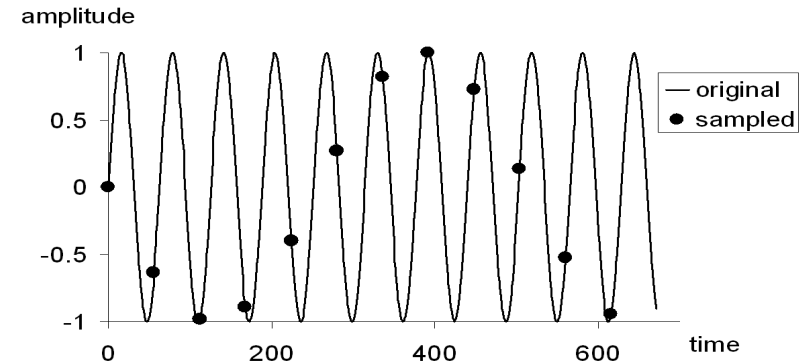
2 bits



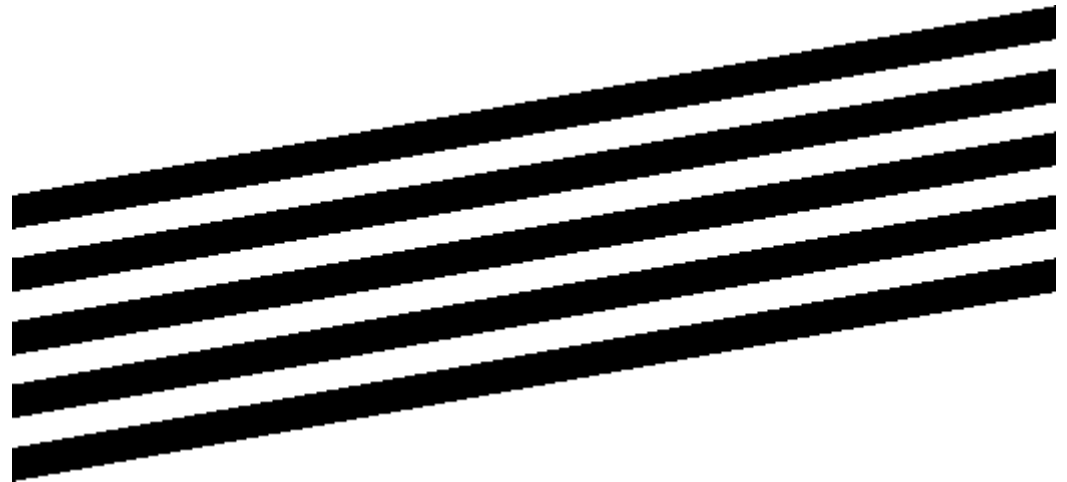
1 bit

I-3 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 :

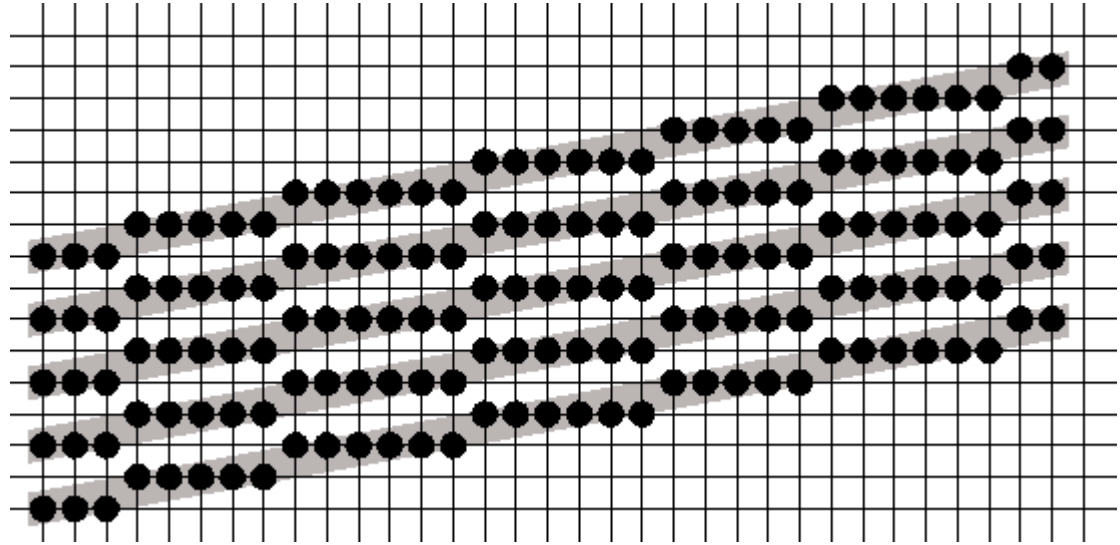


Ce phénomène appelé *aliasing* est encore pire en 2d, car il affecte *la fréquence et la direction* des structures périodiques. Imaginons par exemple qu'on souhaite échantillonner l'image correspondant aux bandes noires ci-contre :

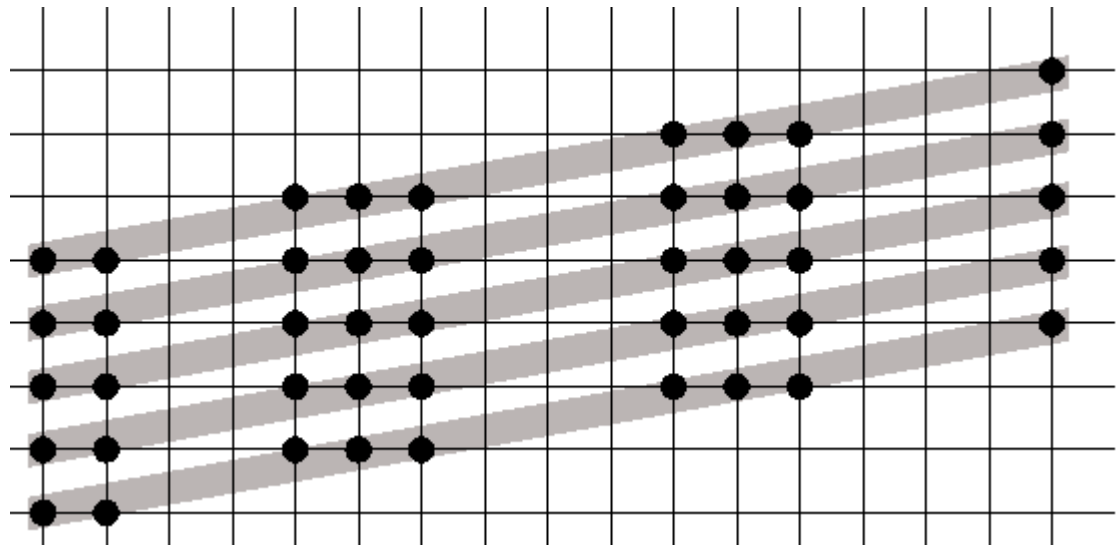


I-3 Echantillonnage et information

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

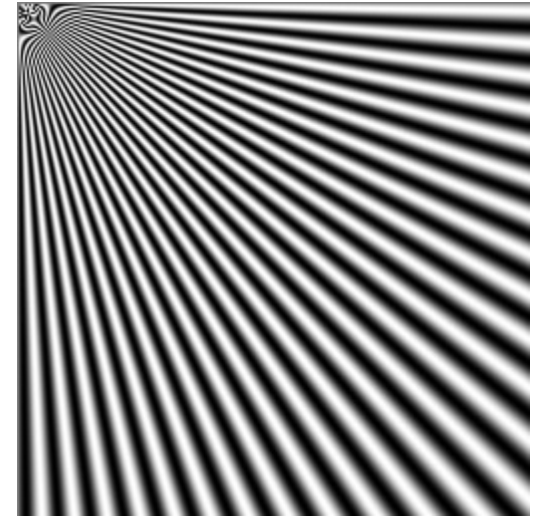


Mais en considérant seulement 1 échantillon sur 2, une structure différente apparaît, dont l'analyse (ici des bandes verticales, plus épaisses) ne sera pas conforme à la réalité de l'objet :



I-3 Echantillonnage et information

Un exemple, sur une image de synthèse :



Et sur une image naturelle :



Image originale



Image sous-échantillonnée

I-4 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.

I-4 Entropie 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, l'*histogramme*).



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

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.

I-3 Outils fondamentaux

a) 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.

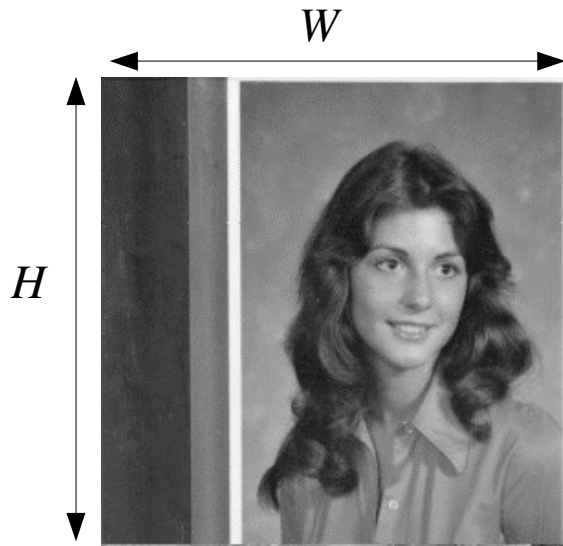
b) La convolution

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.

c) La transformée de Fourier

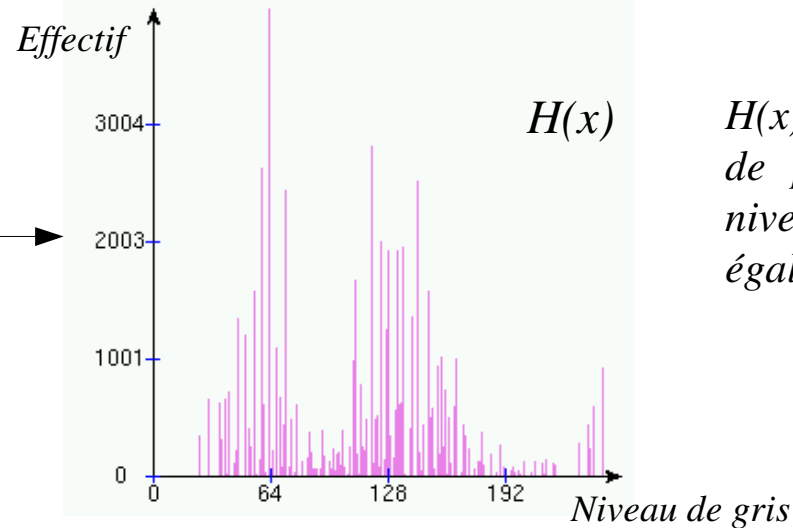
Outil fondamental d'analyse en traitement du signal, le pendant bidimensionnel de la TF et sa version discrète peut être appliqué avec profit aux images numériques. Si son utilisation en tant qu'outil analytique et algorithmique s'est estompée en traitement d'images au profit d'approches mieux adaptées à la localisation spatiale des fréquences (ondelettes), elle reste un outil théorique et pédagogique intéressant.

I-3-a Histogramme



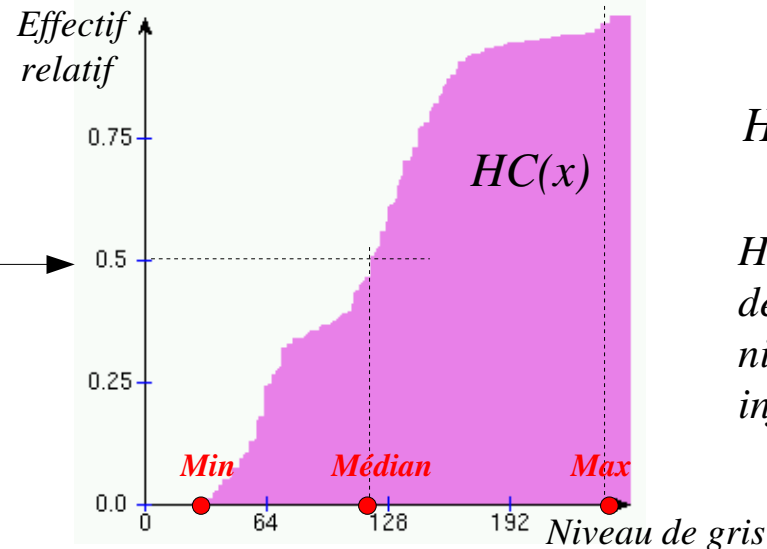
L'histogramme représente la répartition des pixels en fonction de leur niveau de gris. Il fournit diverses informations comme les statistiques d'ordre (voir ci-contre) et peut permettre d'isoler des objets.

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 .

I-3-b Convolution

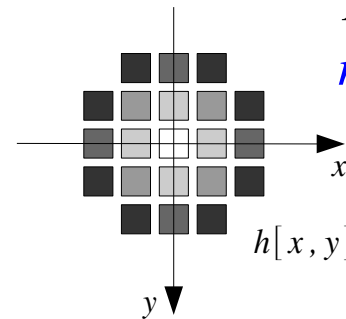
Soit I une image numérique.

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

I-3-c Transformée de Fourier (1)

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

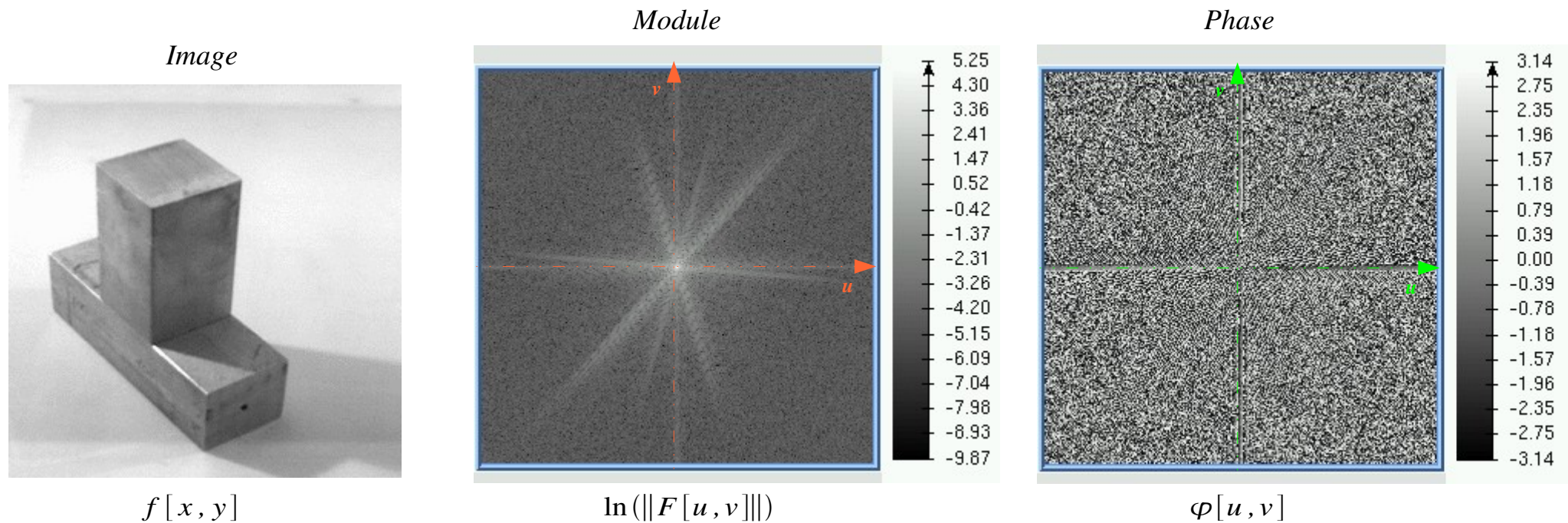
$$F[u, v] = F[u + w, v + h]$$

SYMÉTRIE

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

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

I-3-c Transformée de Fourier (2)



Propriétés de la transformée de Fourier (2) :

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

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

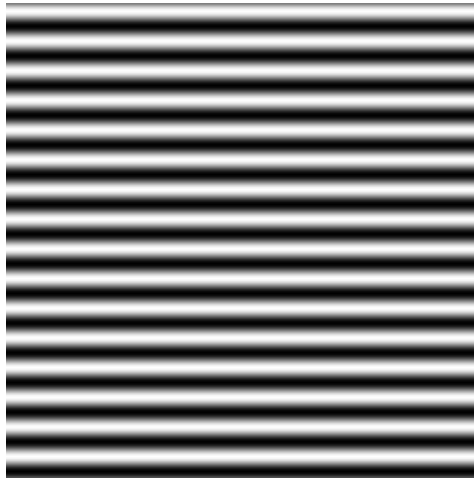
$$\begin{aligned} f[x - x', y - y'] &\rightarrow F[u, v] \cdot e^{-2i\pi(ux' + vy')/wh} \\ f[x, y] \cdot e^{2i\pi(u'x + v'y)/wh} &\rightarrow F[u - u', v - v'] \end{aligned}$$

THÉORÈME DE PARSEVAL

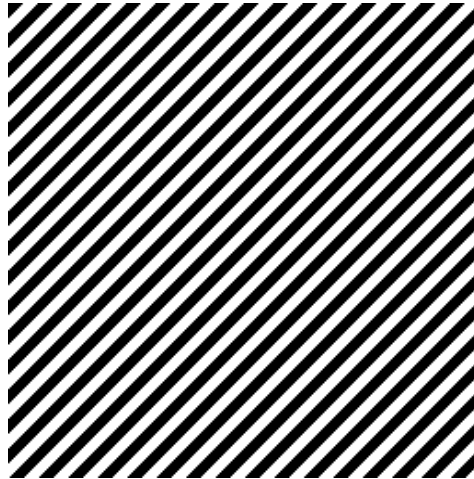
$$\sum_{x=0}^{w-1} \sum_{y=0}^{h-1} \|f[x, y]\|^2 = \frac{1}{wh} \sum_{u=0}^{w-1} \sum_{v=0}^{h-1} \|F[u, v]\|^2$$

QUIZZ - Transformée de Fourier

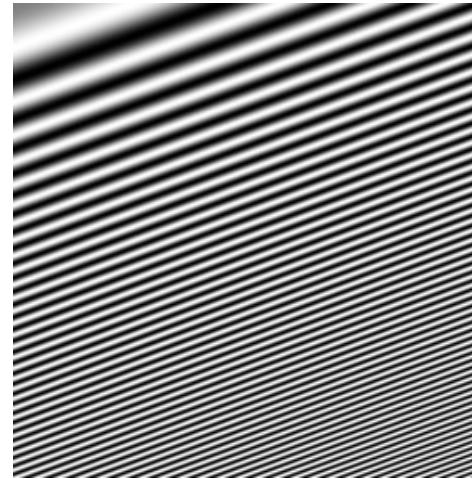
Attribuez à chaque image son spectre de Fourier



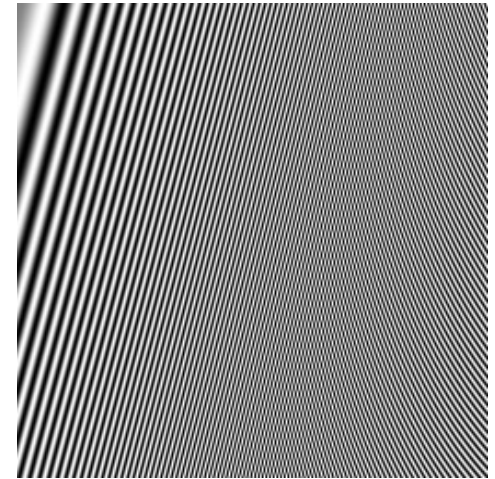
(1)



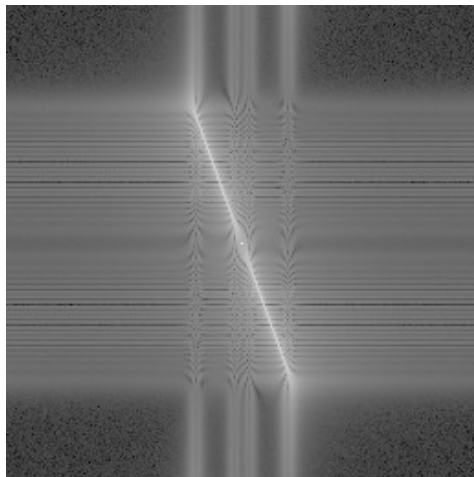
(2)



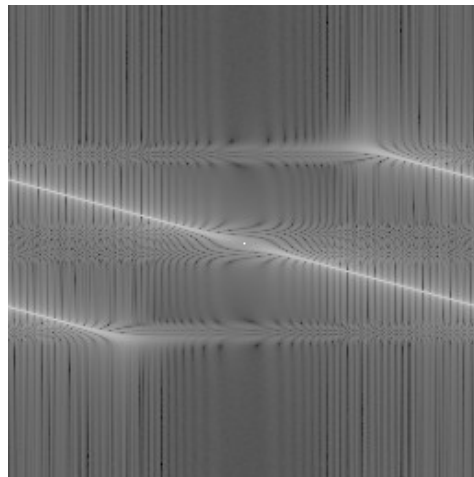
(3)



(4)



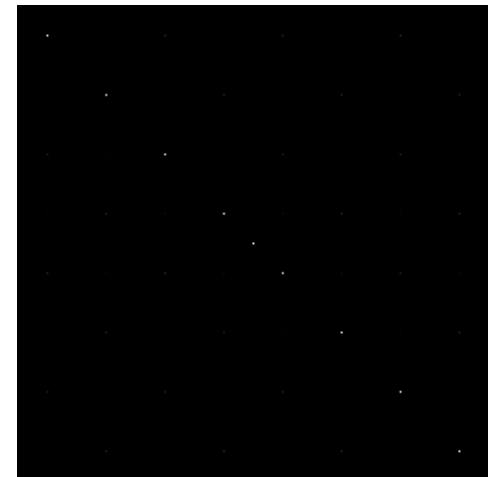
(a)



(b)



(c)



(d)

QUIZZ - Transformée de Fourier

Attribuez à chaque image son spectre de Fourier



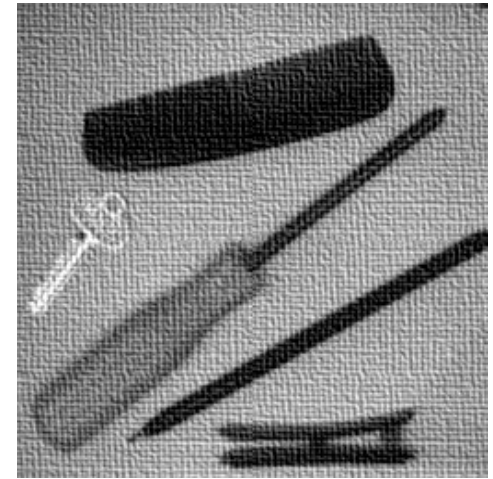
(1)



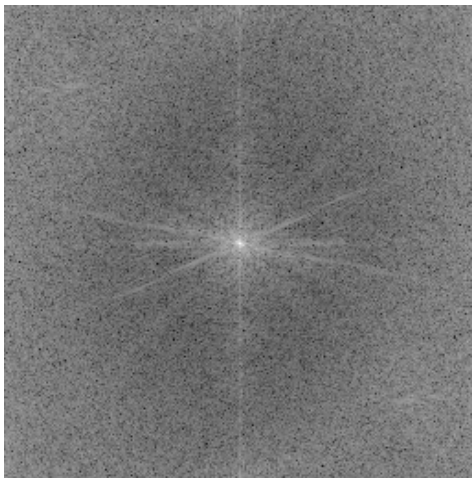
(2)



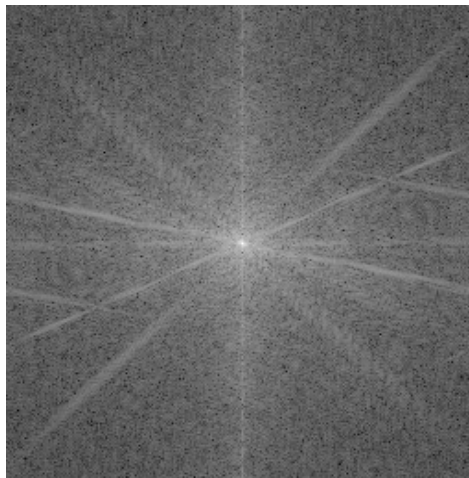
(3)



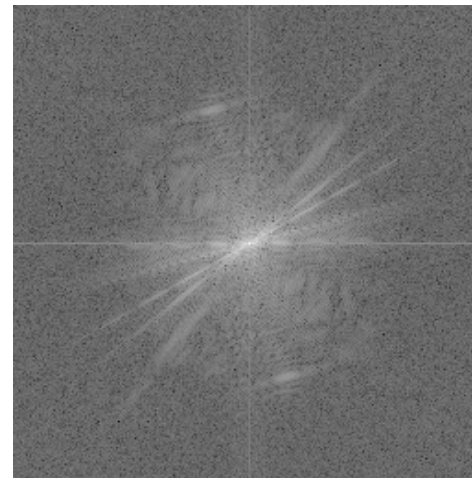
(4)



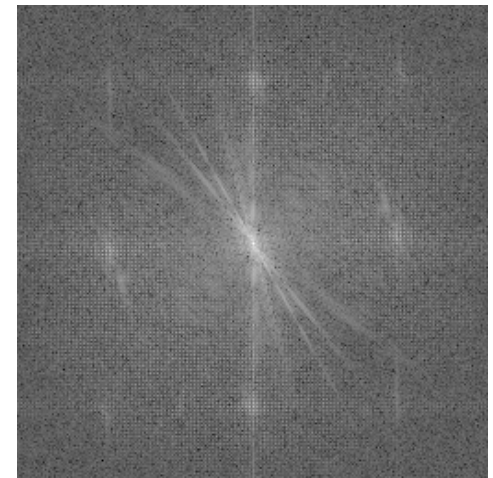
(a)



(b)



(c)



(d)

II 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.

II-1 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.

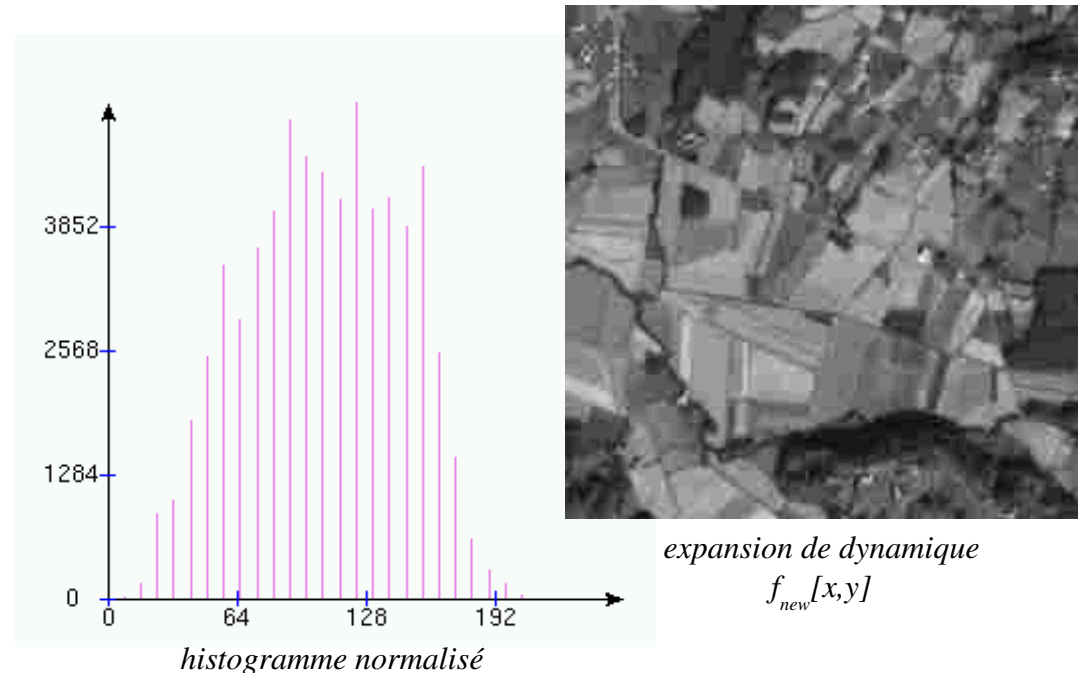
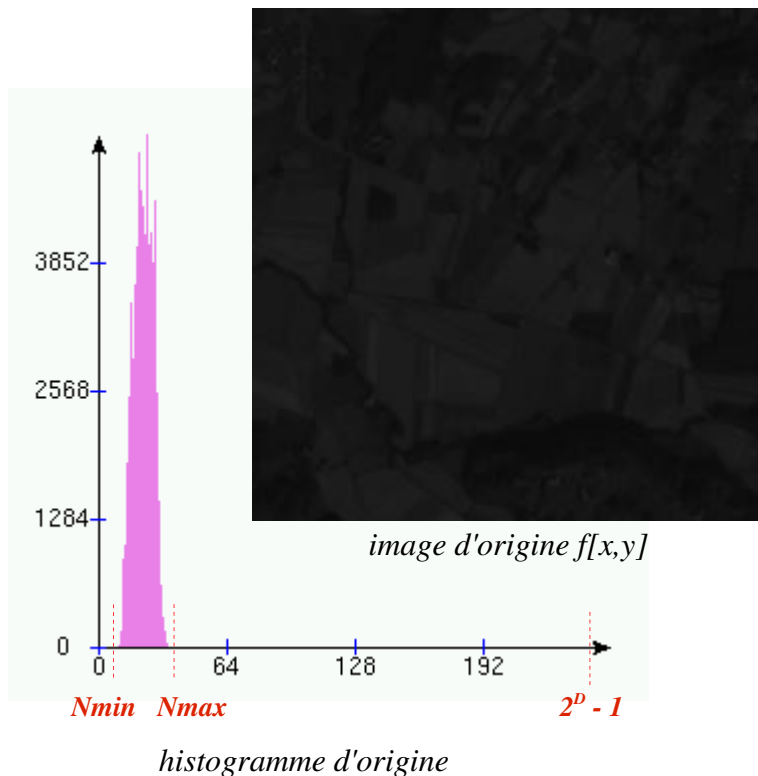
- D : dynamique
- $Nmin$: la plus petite valeur dans l'image
- $Nmax$: la plus grande valeur dans l'image

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

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

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

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



II-2 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*).

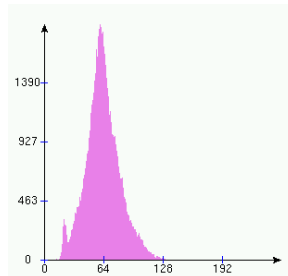
• La technique classique consiste à rendre « le plus linéaire possible » l'histogramme cumulé de l'image en utilisant la transformation suivante :

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

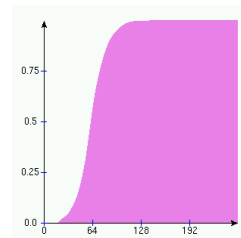
- D : dynamique
- (w, h) : dimension de l'image
- $HC(.)$: histogramme cumulé



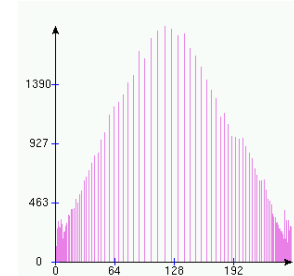
Original $f[x, y]$



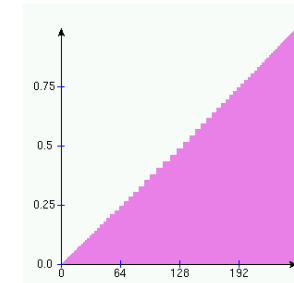
Histogramme de f



Histogramme 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'imager infra-rouge.

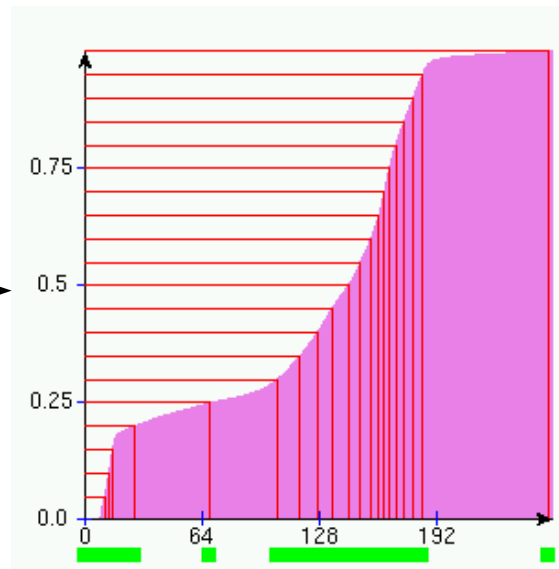
II-3 Histogramme : segmentation

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 Filtres de lissage

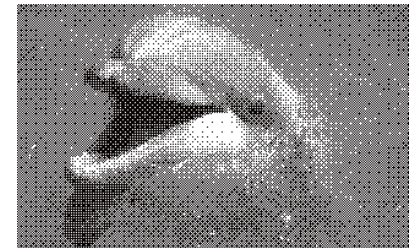
Les filtres de lissage sont des opérateurs qui *éliminent* des éléments *perturbateurs / non significatifs* dans les images numériques, soit pour *améliorer* leur visualisation, soit pour les *simplifier* en but d'un traitement postérieur :



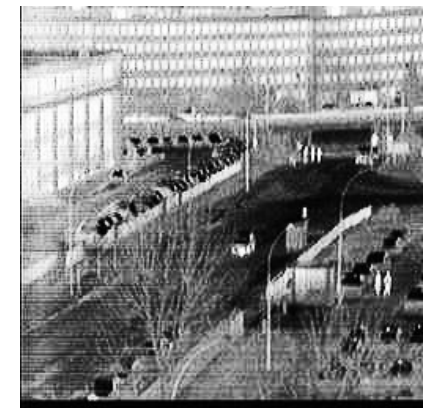
bruit d'acquisition, de numérisation, de transmission : les incertitudes dans les différentes étapes de formation de l'image numérique induisent des fluctuations aléatoires de la valeur des pixels (à droite, bruit gaussien). Les erreurs de transmission font apparaître des valeurs aberrantes (à gauche, bruit impulsif).



bruit de compression : les techniques de compression d'image avec perte produisent une distortion dans l'image, comme cet effet de bloc dans la transformée Jpeg (taux de compression 1/25).



rendu : les images codées en demi-teintes de l'imprimerie présentent à grande échelle un effet pointilliste.



bruit spatial fixe : la non-uniformité des détecteurs dans la matrice de cet imageur infra-rouge entraîne une texturation de l'image.

- PLAN DU CHAPITRE :
- (1) Filtrage dans le domaine de Fourier
 - (2) Filtrage par convolution
 - (3) Implantation des filtres linéaires
 - (4) Filtres non linéaires

III-1 Filtrage dans le domaine de Fourier (1)

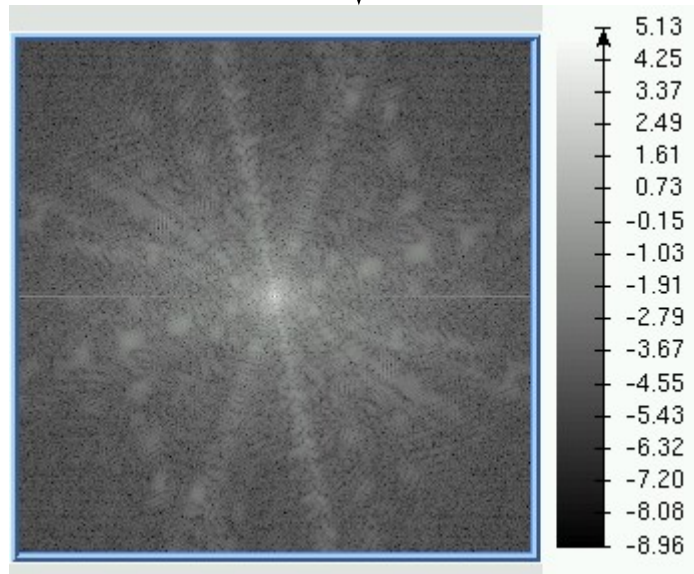
Filtrage passe-bas

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

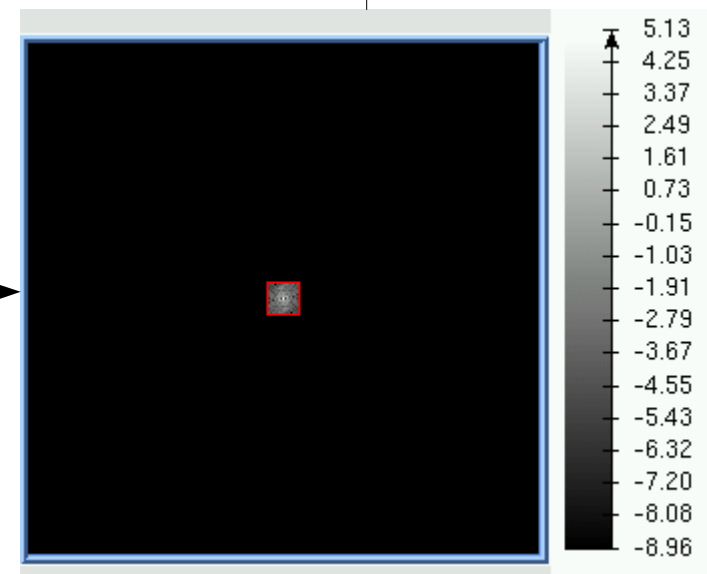


TF

TF^{-1}



Passe-bas



III-1 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 :

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

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

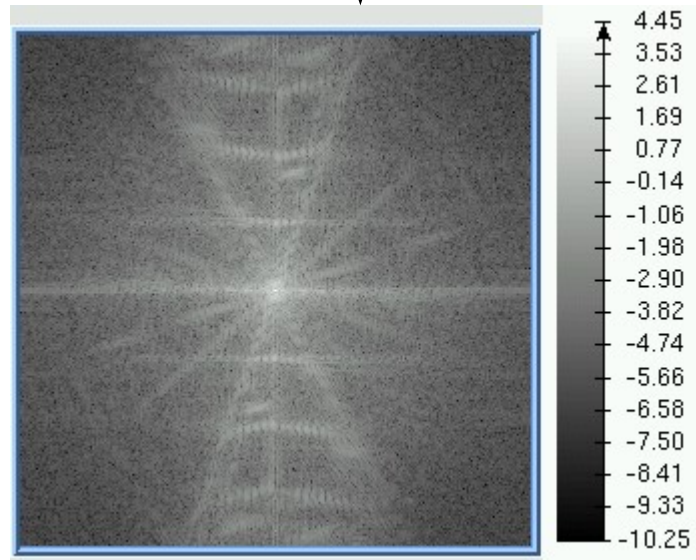
$$F[0,0] = \sum_{x=0}^w \sum_{y=0}^h f[x, y]$$

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

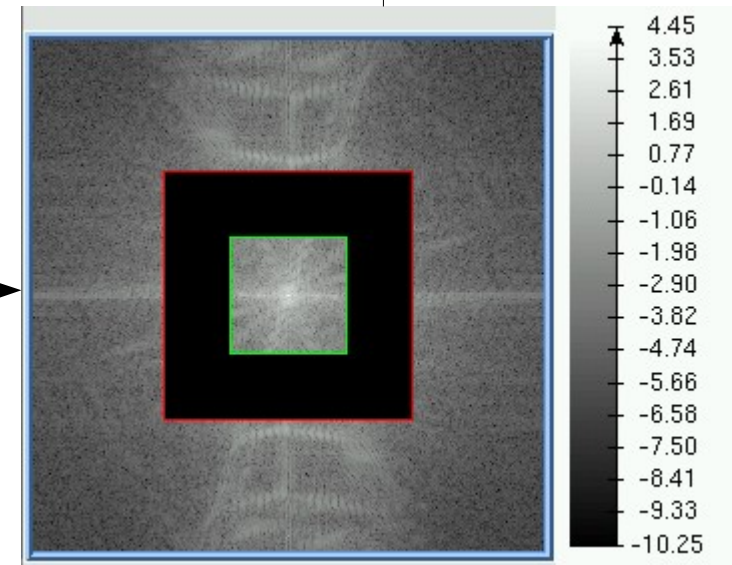


TF

TF⁻¹



Coupe-bande



III-2 Filtrage par convolution (1)

La multiplication dans le domaine fréquentiel correspond à la convolution dans le domaine spatial. 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 famille de filtres parmi les plus utilisés :

Moyenne

** Réponse impulsionnelle :*

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

$$h(x, y) = 0 \text{ sinon}$$

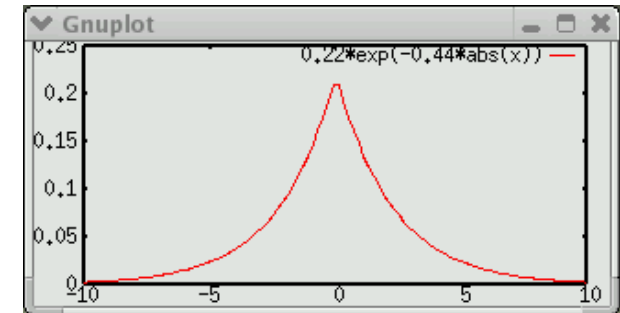
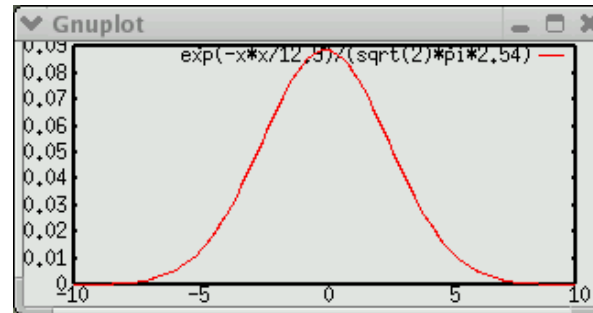
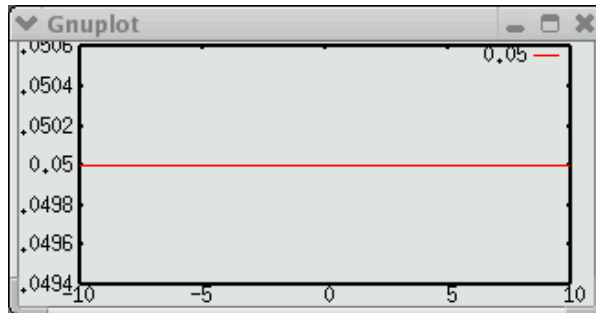
Gauss

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

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 moyeneur (5x5)

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

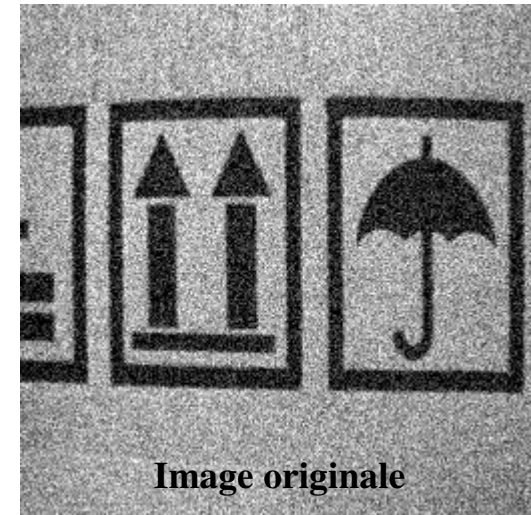
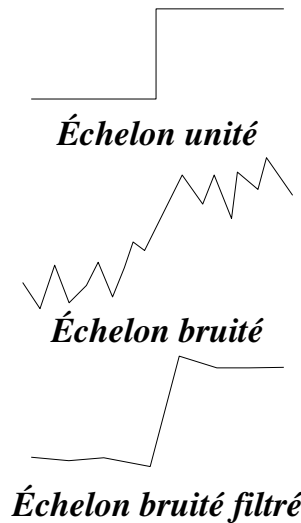
Filtre gaussien ($\sigma = 1,41$)

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

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

III-2 Filtrage par convolution (2)

Coefficient de dispersion : pour un échelon unitaire perturbé par un bruit blanc de variance v^2 , la variance du bruit filtré devient δv^2 .

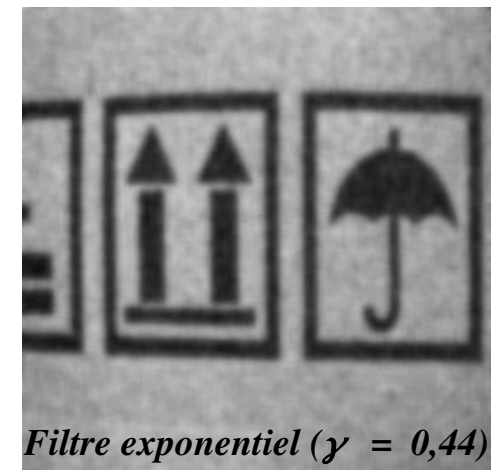
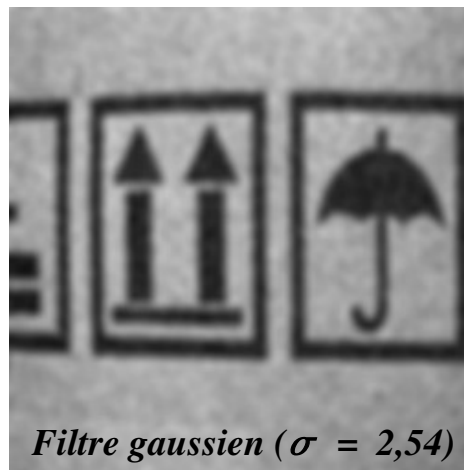
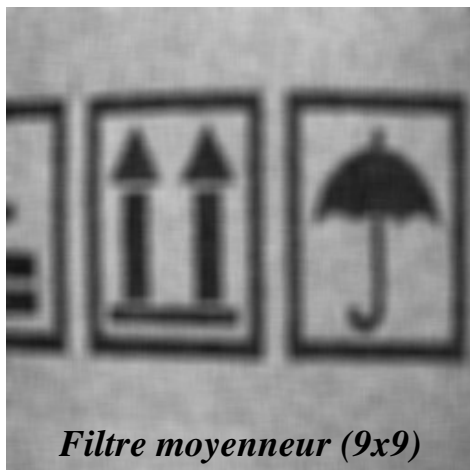


* *Coefficients de dispersion* :

$$\delta = \frac{1}{\lambda^2}$$

$$\delta = \frac{1}{4\pi\sigma^2}$$

$$\delta = \frac{\gamma^2}{16}$$



III-3 Implantation des filtres linéaires

En traitement d'images, les volumes de données traités sont bien sûr très importants. La prise en compte du temps de calcul reste un élément majeur dans les algorithmes en dépit des progrès technologiques exponentiels des microprocesseurs. L'implantation des filtres linéaires, en particulier ceux dont le support est grand, voire infini, est un problème incontournable.

- (a) multiplication dans le domaine de Fourier
- (b) convolution directe par noyau (tronqué)
- (c) noyaux séparables
- (d) implantation récursive des filtres à réponse impulsionnelle infinie

III-3-a Multiplication dans le domaine de Fourier

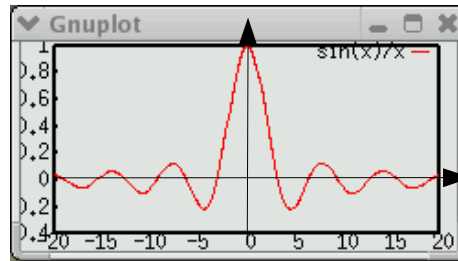
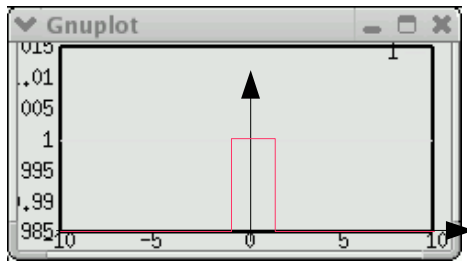
Grâce à la correspondance convolution-produit dans la transformée de Fourier (TF), la convolution de l'image f par un filtre de réponse impulsionnelle h peut se calculer comme la TF inverse du produit $F \cdot H$, où F (resp. H) est la TF de f (resp. h).

CORRESPONDANCE CONVOLUTION / PRODUIT

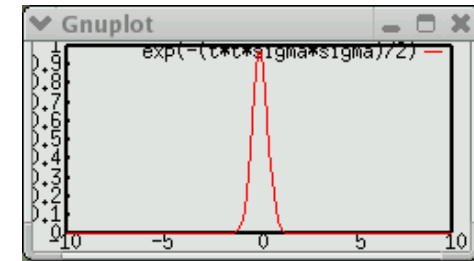
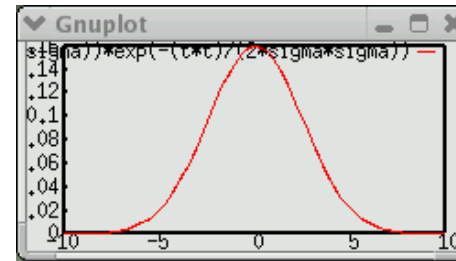
$$f_1[x, y] * f_2[x, y] \rightarrow F_1[u, v] \cdot F_2[u, v]$$

$$f_1[x, y] \cdot f_2[x, y] \rightarrow F_1[u, v] * F_2[u, v]$$

Fonction porte \leftrightarrow *Sinus cardinal*



Gaussienne (σ) \leftrightarrow *Gaussienne ($1/\sigma$)*



La complexité de l'implantation par multiplication dans le domaine fréquentiel est celle de 2 calculs de TF (1 direct + 1 inverse), plus 1 multiplication. Pour une image de taille $N \times N$, le coût de la multiplication est en $O(N^2)$, et en utilisant la transformée de Fourier rapide (FFT), le coût de la TF est en $O(N \cdot \log_2(N))$.

Dans ce cas, la complexité est indépendante de la taille $K \times K$ du noyau de convolution. Ce type d'implantation peut être intéressant pour des gros noyaux, ($K^2 \gg \log_2(N)$). Il nécessite cependant une grande précision dans les valeurs de la TF (représentation en complexes flottants).

III-3-b/c convolution directe / noyaux séparables

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

La convolution de l'image f par un filtre de réponse impulsionnelle h représenté par un noyau fini (éventuellement tronqué) peut être calculé directement par balayage des pixels de f et calcul de la somme des valeurs des voisins de chaque pixel pondérées par les valeurs du noyau de convolution.

La complexité de l'implantation directe pour une image de taille $N \times N$ et pour un noyau de convolution de taille $K \times K$, est en $O(K^2 N^2)$. Le coût par pixel est donc quadratique en fonction du rayon du noyau.

Filtres séparables :

Lorsque la matrice de convolution peut s'écrire comme produit d'un vecteur colonne et d'un vecteur ligne :

$$\llbracket h \rrbracket = \llbracket h_{col} \rrbracket \cdot \llbracket h_{lig} \rrbracket^t$$

Alors : $h[x, y] = h_{col}[x] \cdot h_{lig}[y]$

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

La complexité de l'implantation pour une image de taille $N \times N$ et pour un noyau de convolution de taille $K \times K$, devient $O(KN^2)$. Le coût par pixel est donc linéaire en fonction du rayon du noyau.

Les filtres moyenneur, gaussien, exponentiel sont des filtres séparables.

III-3-d Implantation récursives des filtres IIR

La convolution directe par noyau fini permet d'implanter les filtres à réponse impulsionnelle finie (FIR), mais pose problème dans le cas des filtres à réponse impulsionnelle infinie (IIR). On peut approximer les filtres IIR en tronquant le noyau de convolution (on choisit par exemple des supports de rayon 2σ ou 3σ pour approximer la gaussienne par un filtre FIR). On retiendra cependant que la TF d'un filtre FIR étant à support infini, on ne peut pas éliminer totalement les hautes fréquences avec un filtre FIR.

Certains filtres IIR possèdent la propriété de pouvoir être calculés *de manière récursive*. C'est le cas du filtre exponentiel (cf TP 1), ou de certaines approximation du noyau gaussien. Le filtrage est en général obtenu par un filtre causal, calculé par balayage direct, suivi d'un filtre anti-causal, calculé par un balayage rétrograde :

Ex : filtre IIR 1D horizontal :

$$f[i] = \alpha_0 f[i] + \alpha_1 f[i-1] + \alpha_2 f[i-2] \quad \text{Séquence causale (directe)}$$

$$f[i] = \gamma_0 f[i] + \gamma_1 f[i+1] + \gamma_2 f[i+2] \quad \text{Séquence anti-causale (rétrograde)}$$

La complexité de cette implantation est en $O(N^2)$, elle est en général indépendante des paramètres du noyau de convolution. Elle a de plus donné lieu à des implantations matérielles (circuits spécialisés). Cependant les problèmes de précision nécessitent en général un passage en nombre flottant et donc une augmentation de la dynamique.

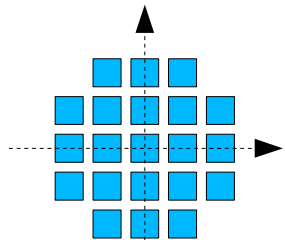
III-4 Filtres non linéaires

- Deux aspects du lissage sont concernés par le filtrage non linéaire :
- *Le bruit impulsionnel* : les filtres linéaires éliminent mal les valeurs aberrantes.
- *L'intégrité des frontières* : on souhaiterait éliminer le bruit sans rendre flous les frontières des objets.

- (a) Filtres d'ordre, médian
- (b) Filtres non linéaires divers – ex : Nagao
- (c) NL-Means
- (d) Filtres morphologiques

III-4-a Filtres d'ordre, médian

Les filtres d'ordres procèdent en remplaçant les valeurs de chaque pixel par la valeur qui occupe *un certain rang* lorsqu'on trie les valeurs observées dans *un certain voisinage* du pixel.



voisinage : élément
structurant

les valeurs dans le voisinage de (x,y) : $V(x,y) = \{a_1, a_2, \dots, a_N\}$

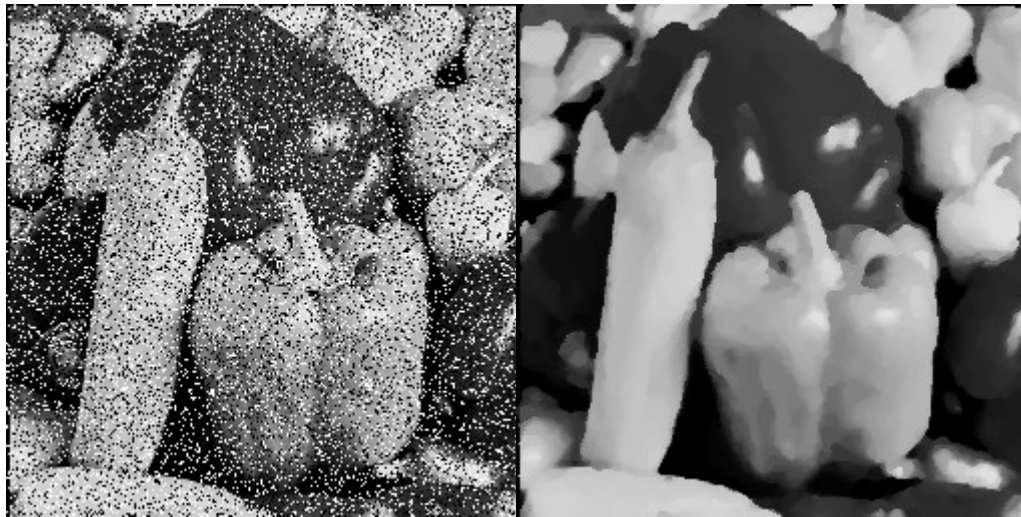
soit $\{b_1, b_2, \dots, b_N\}$ permutation de $\{a_1, a_2, \dots, a_N\}$ telle que $b_1 \leq b_2 \leq \dots \leq b_N$

alors le filtre d'ordre de rang k est défini par : $\rho_k[x, y] = b_k$

pour $k=N/2$, on parle de *filtre médian*, pour $k=1$, d'*érosion morphologique*, pour $k=N$, de *dilatation morphologique*.

Implantations du médian :

- calcul d'histogrammes locaux
- tri des valeurs dans le voisinage (Quick Sort)
- tri incrémental
- .../...



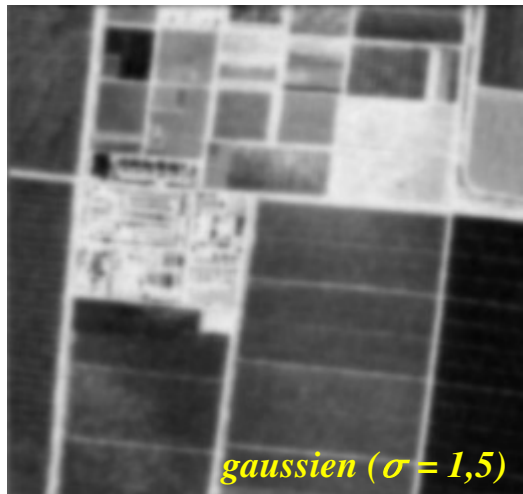
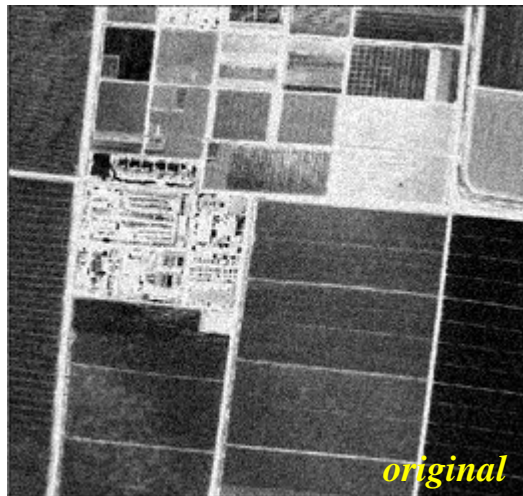
ex : bruit impulsionnel traité par un filtre médian
(voisinage comme ci-dessus).



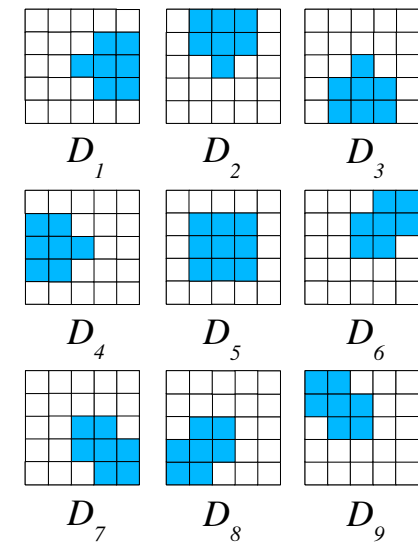
opérateurs morphologiques : à gauche Original au centre Érosion
à droite Dilatation (élément structurant comme ci-dessus)

III-4-b Filtres non linéaires divers

On trouve dans la littérature de nombreux filtres combinant *filtres d'ordre*, *moyennes robustes* (opérations linéaires éliminant les valeurs marginales), et *anisotropie* (le support des opérations s'adapte en fonction des frontières locales). Nous décrivons ici comme exemple le *filtre de Nagao*.



Le filtre de Nagao examine la fenêtre 5x5 centrée sur chaque pixel. 9 domaines sont définis dans cette fenêtre (voir figure). On calcule pour chaque domaine D_i la moyenne μ_i et la variance v_i . Le résultat de l'opérateur est la moyenne du domaine qui présente la plus faible variance.



Les 9 fenêtres de Nagao

III-4-c NL-Means

Le filtre NL-Means (Buades et al 2005), est parmi les meilleurs méthodes de débruitage aujourd'hui en termes de qualité. Il s'agit d'une moyenne pondérée calculée en chaque pixel (comme dans une convolution), mais où le poids attribué à chaque pixel ne dépend pas de la distance entre les pixels, mais de leur similarité vis-à-vis de l'image traitée.

Formulation générale de la moyenne pondérée :

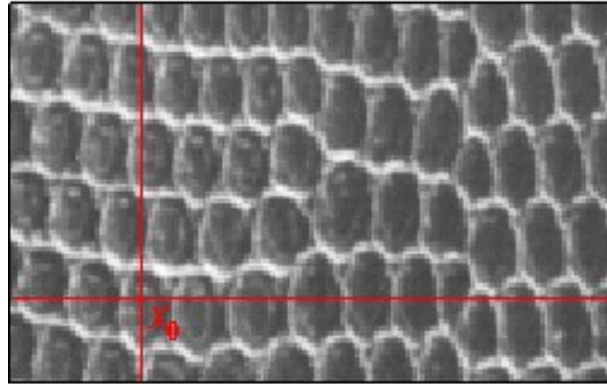
$$f_{new}(p) = \frac{1}{\pi(p)} \sum_{q \in N(p)} w(p, q) f(q)$$

Où :

- f et f_{new} représentent l'image respectivement avant et après le filtrage,
- q et p sont les pixels,
- $N(p)$ représente un « voisinage » de p ,
- $\pi(p)$ la fonction de normalisation :
$$\pi(p) = \sum_{q \in N(p)} w(p, q)$$
- $w(p, q)$ le poids relatif de q par rapport à p .

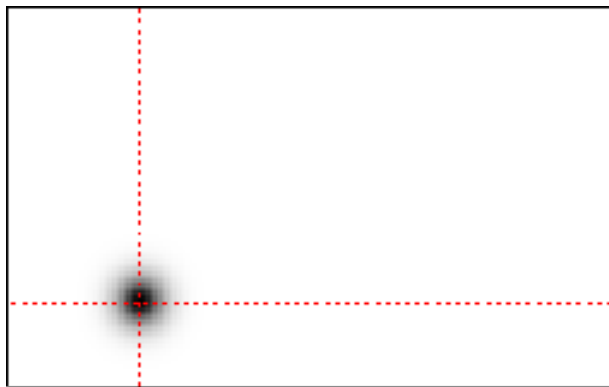
C'est la fonction poids $w(p, q)$ qui différencie les NL-Means d'une convolution classique...

III-4-c NL-Means



Convolution gaussienne

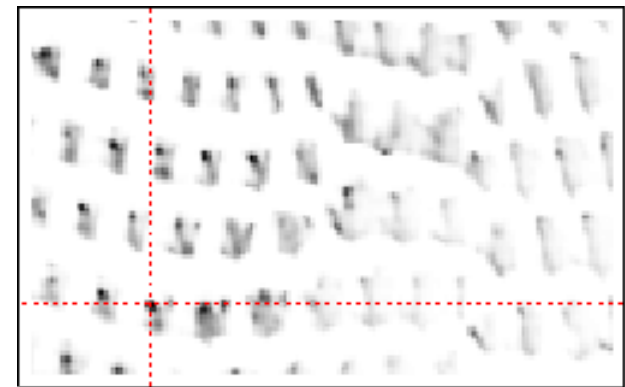
$$\omega(p, q) = e^{\frac{-\|p-q\|^2}{\sigma^2}}$$



Les pixels à poids significatifs sont concentrés autour du pixel p (*Local*).

NL-Means

$$\omega(p, q) = e^{\frac{-d_f(p, q)^2}{h^2}} \quad d_f(p, q) = \sum_k (f(p+k) - f(q+k))^2$$



Les pixels à poids significatifs sont potentiellement partout dans l'image (*Non Local*).

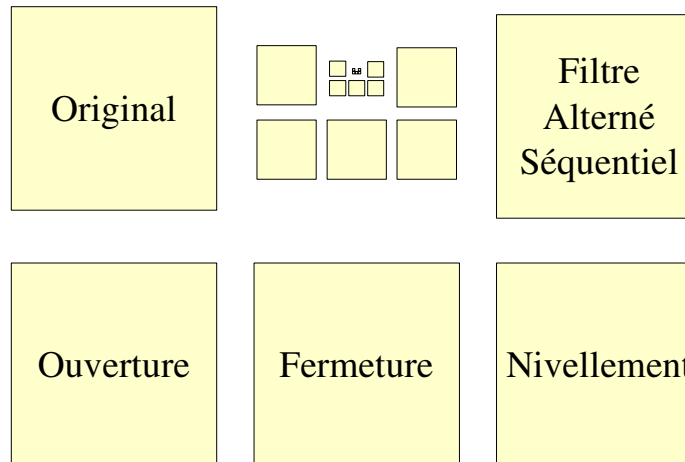
III-4-c NL-Means

Le filtre NL-Means fournit de très bons résultats en exploitant les corrélations dans les textures d'images naturelles.

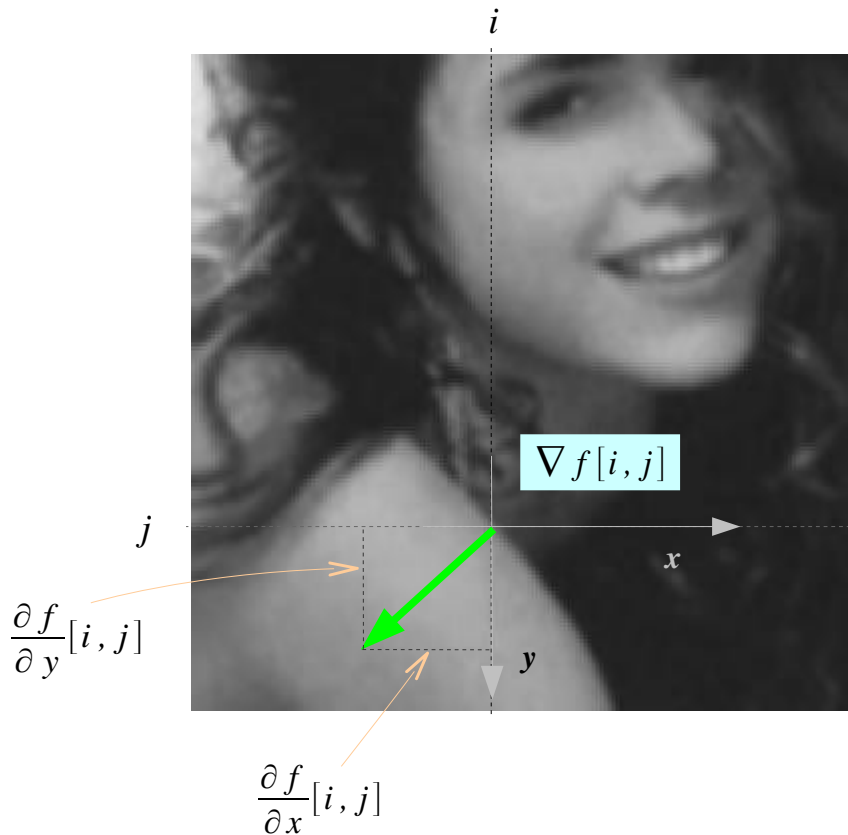


III-4-d Filtres morphologiques

...pour mémoire. Voir cours de morphologie mathématique.



IV Filtres dérivateurs



Les *variations locales d'intensité* constituent une source primordiale d'information en traitement d'images. Elles sont mesurées par le *gradient*, fonction *vectorielle* des pixels $[i, j]$:

$$\nabla f[i, j] = \left(\frac{\partial f}{\partial x}[i, j], \frac{\partial f}{\partial y}[i, j] \right)$$

D'autres grandeurs différentielles sont utilisées en traitement d'images, comme le *laplacien*, fonction *scalaire* de $[i, j]$:

$$\Delta f[i, j] = \frac{\partial^2 f}{\partial x^2}[i, j] + \frac{\partial^2 f}{\partial y^2}[i, j]$$

ou encore le *hessien*, fonction *matricielle* de $[i, j]$:

$$H_f[i, j] = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}[i, j] & \frac{\partial^2 f}{\partial x \partial y}[i, j] \\ \frac{\partial^2 f}{\partial x \partial y}[i, j] & \frac{\partial^2 f}{\partial y^2}[i, j] \end{pmatrix}$$

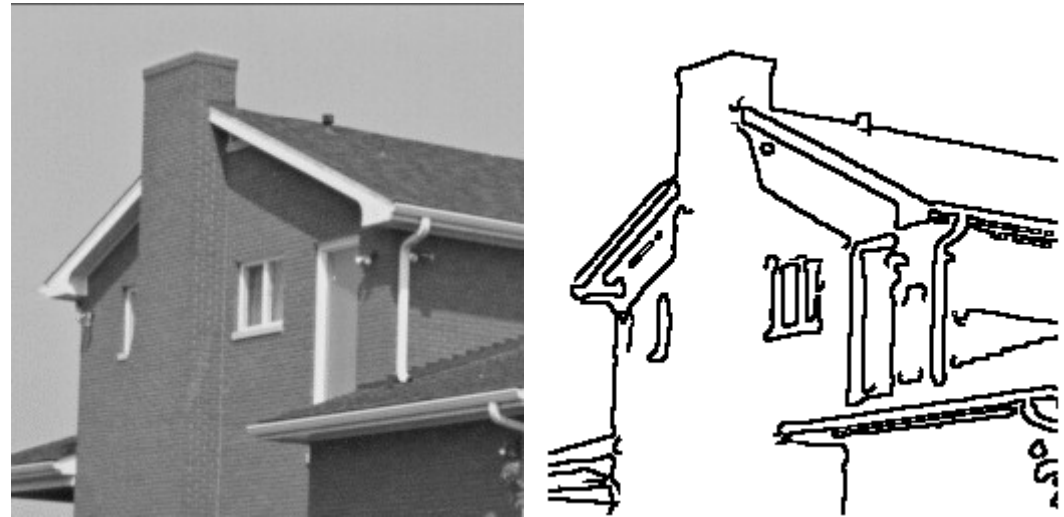
Le problème du calcul des filtres dérivateurs dans les images numériques est l'*approximation* de ces grandeurs différentielles dans notre espace discret ; on s'intéresse aussi à leur *utilisation* : réhaussement, détection de contours,...

IV Filtres dérivateurs

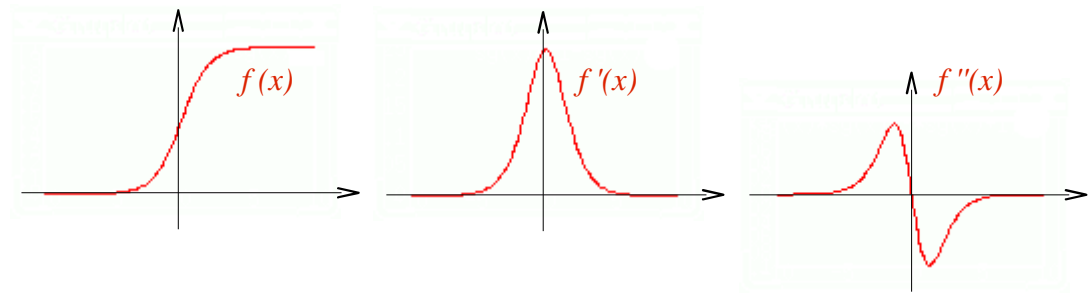
Les filtres dérivateurs sont des opérateurs qui mettent en évidence certaines variations spatiales dans les images. Ils sont utilisés comme traitements de base dans de nombreuses opérations, comme le réhaussement de contraste ou la détection de contours :



réhaussement de contraste obtenu par combinaison linéaire avec le laplacien.



les contours (image de droite) constituent une simplification de l'image utile dans de nombreuses applications. Dans les approches linéaires, ils sont en général obtenus à partir des maxima locaux de la dérivée première, ou des passages par zéro de la dérivée seconde :



PLAN DU CHAPITRE :

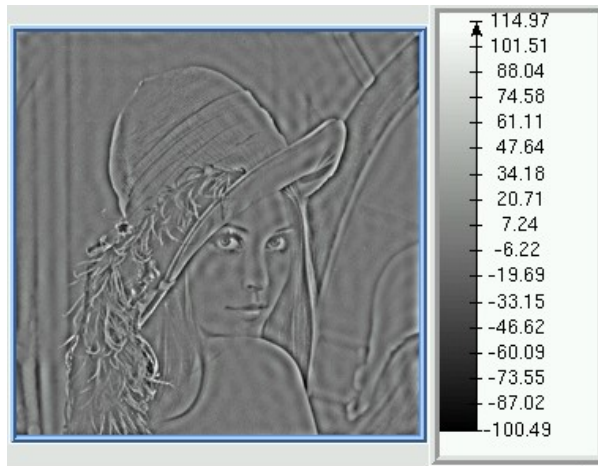
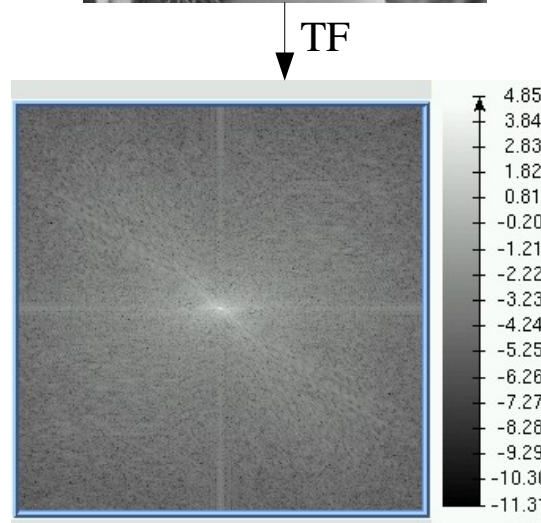
- (1) Filtrage dans le domaine de Fourier
- (2) Filtrage par convolution

IV-1 Filtrage dans le domaine de Fourier

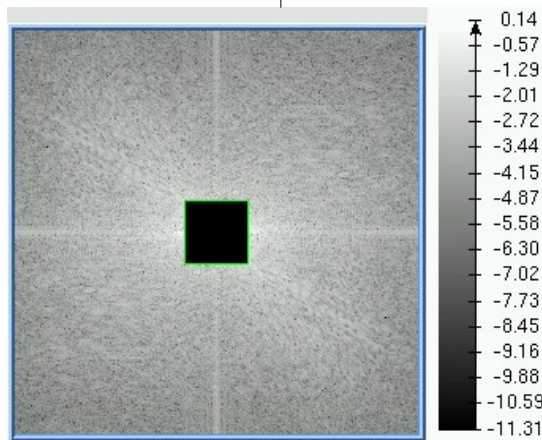
Le filtrage passe-haut correspond à la multiplication dans le domaine fréquentiel d'une fonction porte-complémentaire.

Filtrage passe-haut

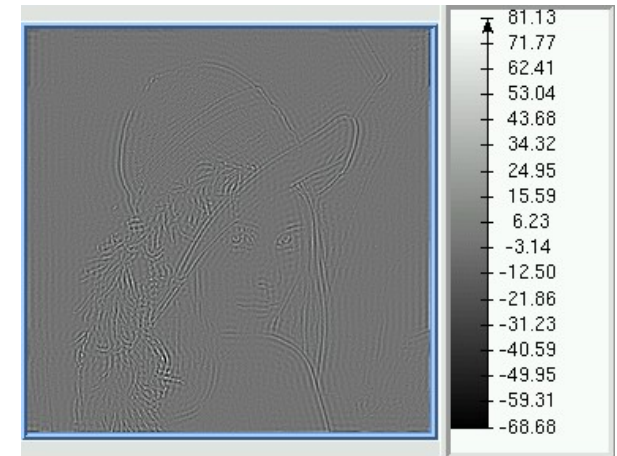
Image originale



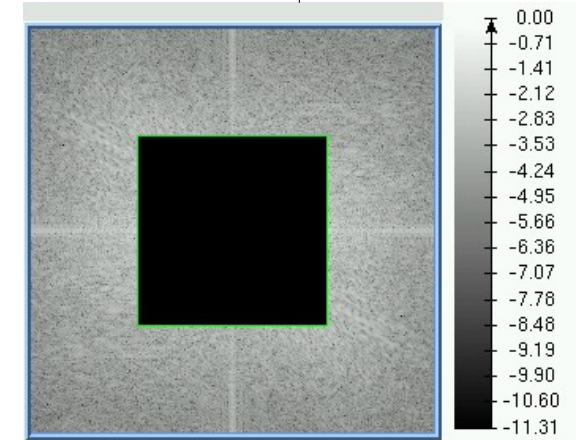
TF^{-1}



Filtre passe-haut $\phi = 20$



TF^{-1}



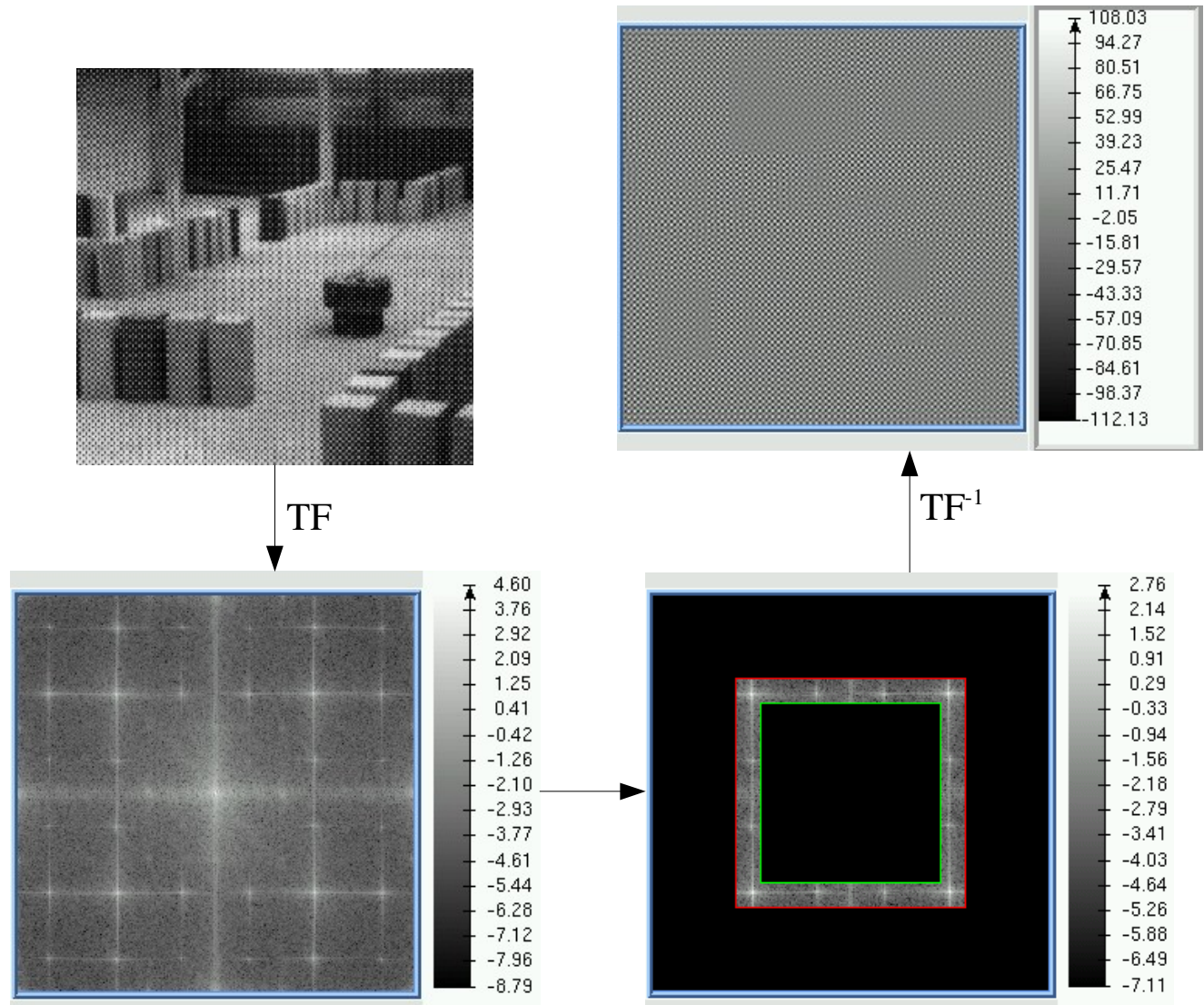
Filtre passe-haut $\phi = 60$

IV-1 Filtrage dans le domaine de Fourier

Filtrage passe-bande

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.



IV-2 Filtrage par convolution

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

Par ex : $\begin{bmatrix} -1 & 1 \end{bmatrix}$, 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 $\begin{bmatrix} -1 & 0 & 1 \end{bmatrix}$, 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) : $\begin{bmatrix} 1 & 2 & 1 \end{bmatrix}$

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

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

On peut ensuite calculer la norme du gradient :

Et son orientation :

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

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

IV-2 Filtrage par convolution



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*

IV-2 Filtrage par convolution

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

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

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 :

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

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



IV-2 TI et portée des opérateurs

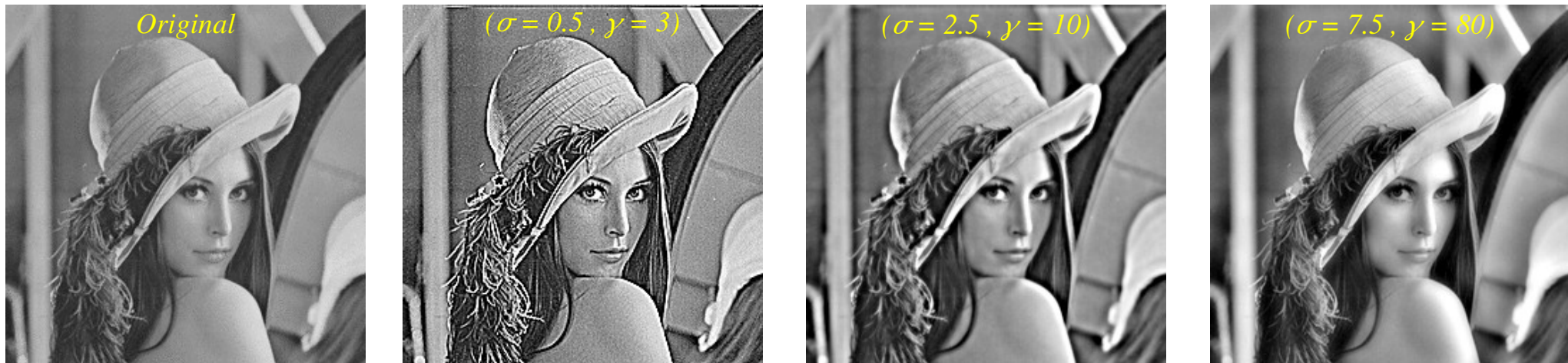
Les opérateurs de traitement d'images possèdent une certaine portée correspondant à l'amplitude du voisinage qui interagit.

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

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

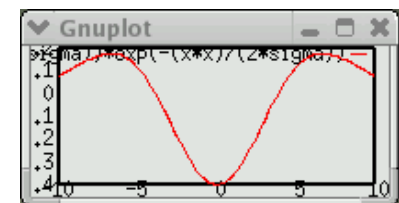
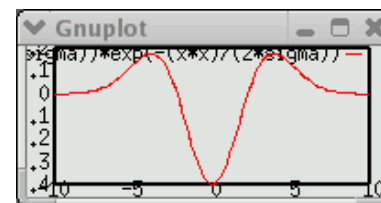
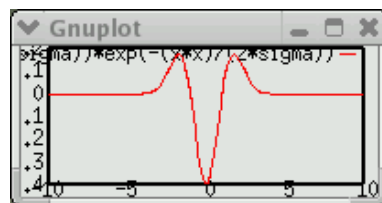
En soustrayant le laplacien à l'image original, on augmente le contraste.

Mais le contraste est une notion multi-échelle :



Il faut donc des mécanismes pour adapter la portée des opérateurs (voir cours espaces d'échelle).

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



Conclusion

A retenir pour ce cours :

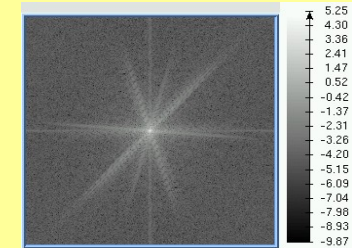
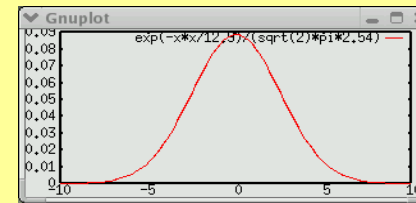
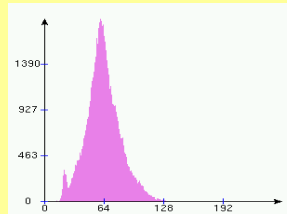
(1) Qu'est-ce qu'une image numérique ?

- échantillonnage
- quantification



(2) Les outils fondamentaux

- histogramme
- convolution
- transformée de Fourier



(3) Filtrage

- Filtres lisseurs
- Filtres différentiateurs



(4) Grandeurs différentielles

- Calcul par différences finies

$$\nabla f[i, j] = \left(\frac{\partial f}{\partial x}[i, j], \frac{\partial f}{\partial y}[i, j] \right)$$