

# *Traitement d'images et vision*

## *Les espaces d'échelle en analyse d'images*

Cours TDI – Master IA&D – Université Paris 6

*Antoine MANZANERA – ENSTA/LEI*

## *Plan du cours « Espaces d'échelles »*

- Intérêt des espaces d'échelle en traitement d'images et vision
  - Pyramides et échantillonnage des images
    - Espaces d'échelles linéaires
    - Exemple 1 : *détection de contours*
  - Exemple 2 : *détection de points anguleux*
    - Pyramides et codages
    - Transformée en ondelettes
  - Exemple 3 : *compression d'images fixes*
    - Espaces d'échelles et diffusion
    - Espaces d'échelles non linéaires
    - Espaces d'échelles morphologiques

## Espaces d'échelles et traitement d'images

La théorie des espaces d'échelle (*Scale space theory*) est un modèle qui est apparu progressivement (79-95-...) dans le domaine de la *vision artificielle*, principalement pour les aspects *bas niveau* (i.e. *traitement du signal*), pour prendre en compte la nature résolument *multi-échelles* des données images.

Les objets réels, la réalité physique, par opposition aux objets mathématiques, existent sous forme d'entités bien différentes selon le niveau d'échelle auquel on les considère:



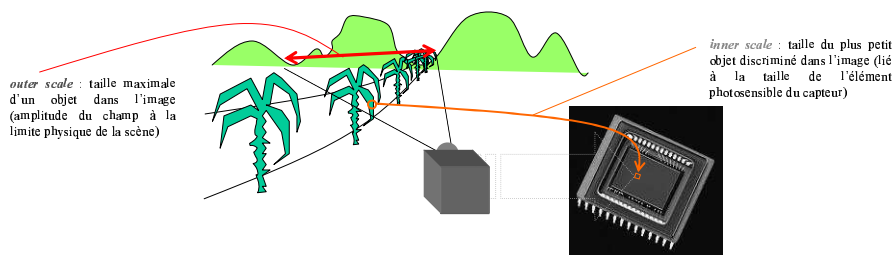
Cours TDI - Master IA&D - UPMC P6

Antoine MANZANERA - ENSTA/LEI

3

## Espaces d'échelles et traitement d'images

Les capteurs d'images utilisés en vision artificielle sont capables d'appréhender un certain domaine dans l'espace des échelles d'une scène réelle, correspondant à l'intervalle entre l'échelle interne (*inner scale*) ou *résolution* du capteur et l'échelle externe (*outer scale*) ou *champ* du capteur.



Le besoin d'une représentation multi-échelles émerge naturellement du développement de méthodes automatiques d'analyse et d'extraction d'information à partir de capteurs. Pour extraire une information d'une donnée image, il est nécessaire de lui appliquer certaines transformations (opérateurs). La nature de l'information alors obtenue est fortement déterminée par le rapport entre la *taille des structures* concernées dans l'image et la *taille des opérateurs* appliqués. Un des problèmes fondamentaux en traitement d'image est : quels opérateurs utiliser, où les appliquer, et quelles doivent être leurs tailles.

[Lindeberg 96]

Cours TDI - Master IA&D - UPMC P6

Antoine MANZANERA - ENSTA/LEI

4

## Introduction : Portée des opérateurs de TI

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

Exemple : l'opérateur de rehaussement de contrastes (unsharp masking) :



$f[x, y]$



$\Delta f[x, y]$



$$R_f[x, y] = f[x, y] - \gamma \cdot \Delta f[x, y]$$

( $\gamma$ : gain)

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

Rappel : le laplacien est défini par  $\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$

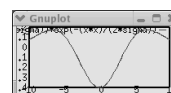
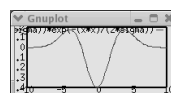
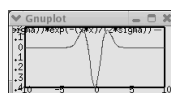
## Introduction : Portée des opérateurs de TI

Mais le contraste est une notion multi-échelle :



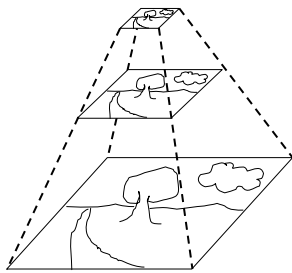
Il faut donc des mécanismes pour adapter la portée des opérateurs.

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



## Espaces d'échelles et traitement d'images

A l'exception de certaines applications pour lesquelles on maîtrise, au moins en partie, l'environnement (contrôle industriel, imagerie microscopique...), les échelles associées à une tâche de vision particulière ne sont pas connues à l'avance. C'est pourquoi il est essentiel que les plus bas niveaux du système de vision (traitements d'images...) considèrent l'image à plusieurs niveaux d'échelles ; c'est à ce titre que les espaces d'échelles constituent un modèle formel adapté.



### Utilisation de la notion de multi-échelle :

- Au niveau *signal* :

représentation (codage) → filtrage → segmentation

- Au niveau *abstrait* :

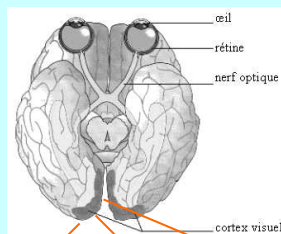
← raisonnement hiérarchique →

### Intérêts algorithmiques du calcul multi-échelle :

- Robustesse
- Réduction du temps de calcul

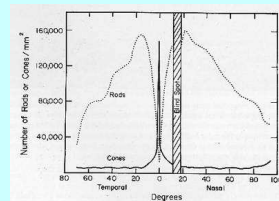
## Liens avec la vision biologique

### « Codage » dans les canaux fréquentiels corticaux



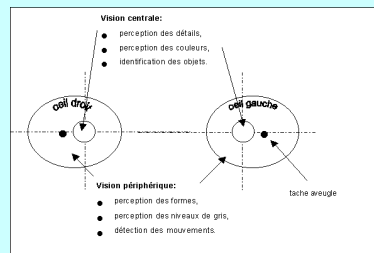
### Vision rétinienne : acuité décroissante du centre à la périphérie

distribution inhomogène des photorécepteurs de la rétine (cônes et bâtonnets) :

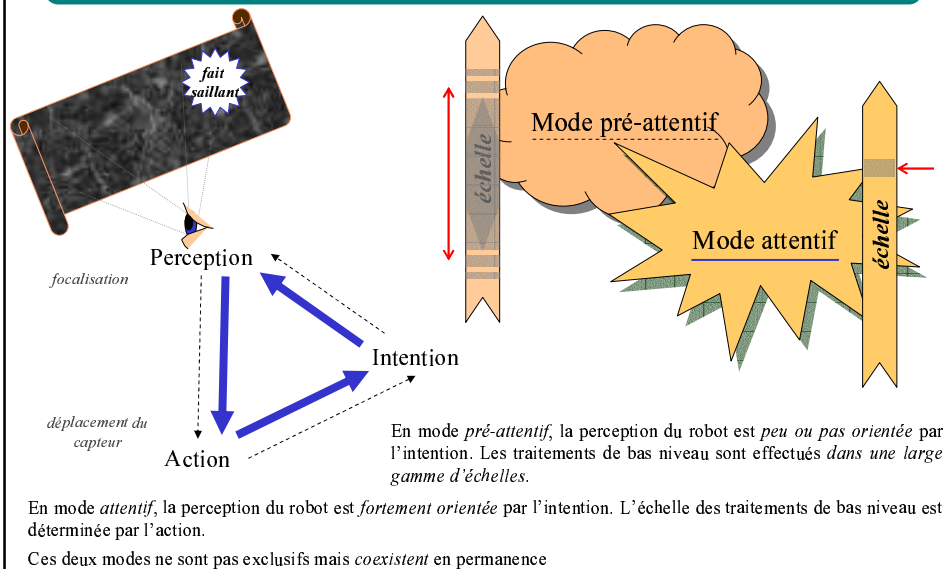


- focus d'attention
- zones préattentives

- détection du mouvement
- tracking
- vision active



## Espaces d'échelles et vision active

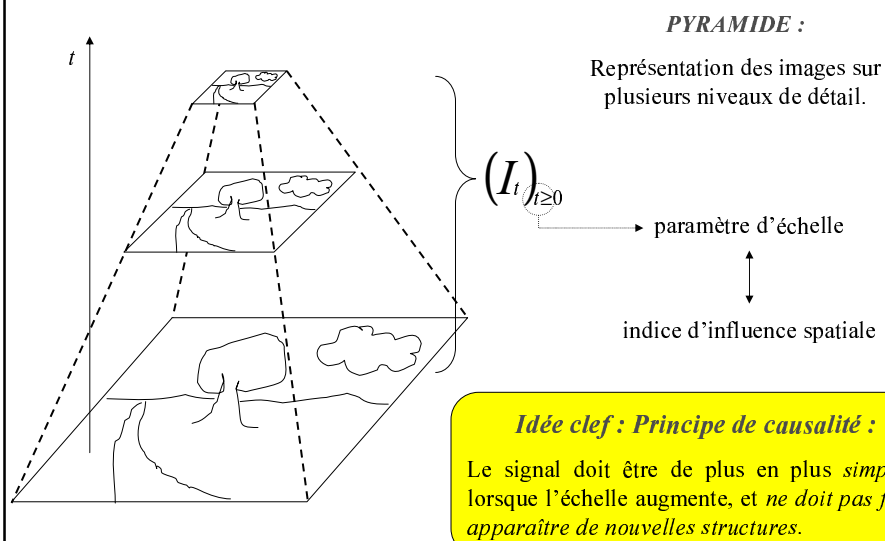


Cours TDI – Master IA&D – UPMC P6

Antoine MANZANERA – ENSTA/LEI

9

## La pyramide visuelle

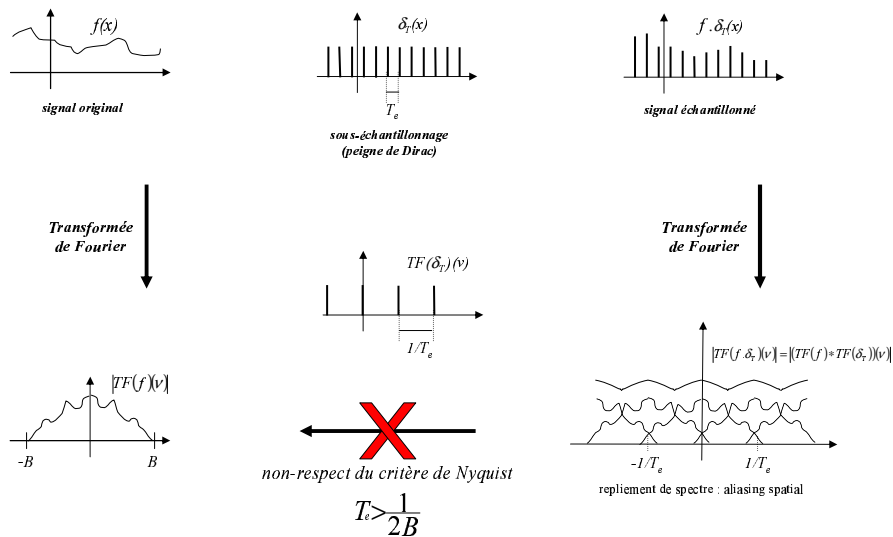


Cours TDI – Master IA&D – UPMC P6

Antoine MANZANERA – ENSTA/LEI

10

## Pyramides et sous-échantillonnage



Cours TDI – Master I&A D – UPMC P6

Antoine MANZANERA – ENSTA/LEI

11

## Pyramides et sous-échantillonnage



Illustration en 2D :

Suite d'images obtenue par quantification



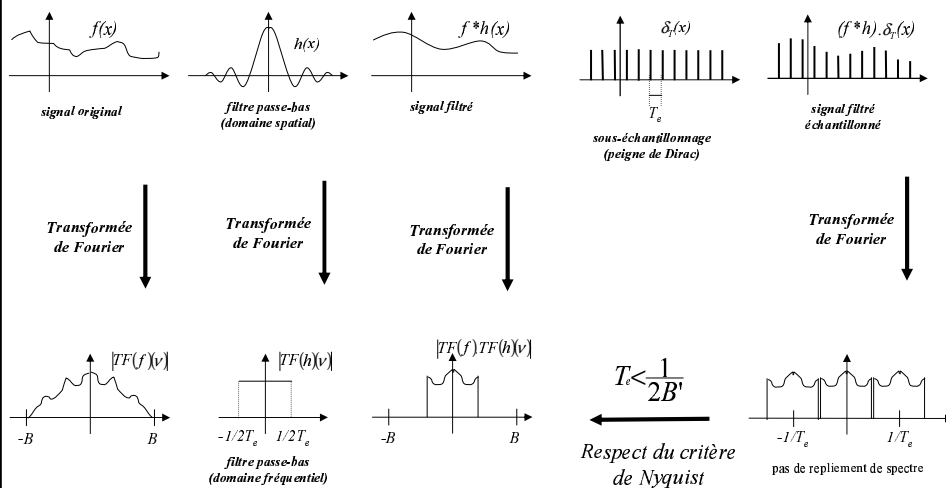
Le non respect du critère de Nyquist introduit un *aliasing spatial* qui correspond au non-respect du principe de causalité de l'espace d'échelle.

Cours TDI – Master I&A D – UPMC P6

Antoine MANZANERA – ENSTA/LEI

12

## Pyramides et sous-échantillonnage



Cours TDI – Master IA&D – UPMC P6

Antoine MANZANERA – ENSTA/LEI

13

## Pyramides et sous-échantillonnage



Illustration en 2D :

**Pyramide de Nyquist**

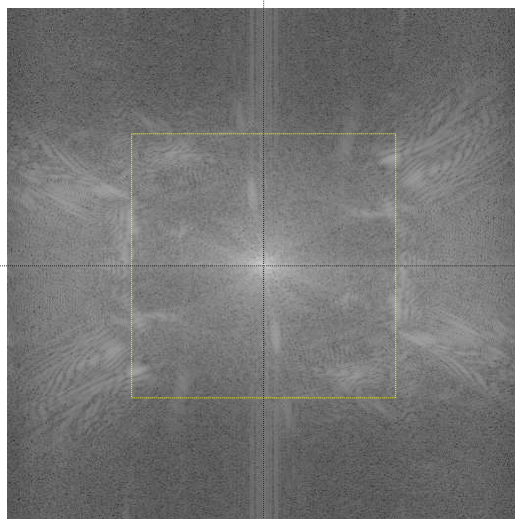


Cours TDI – Master IA&D – UPMC P6

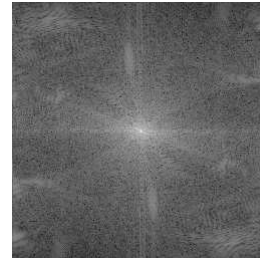
Antoine MANZANERA – ENSTA/LEI

14

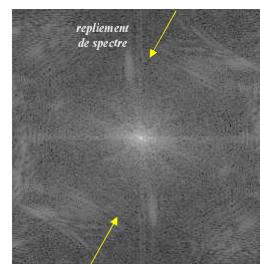
## Pyramides et sous-échantillonnage



Spectre de Fourier de l'image à pleine résolution



Spectre de l'image filtrée et quantifiée



Spectre de l'image simplement quantifiée

Cours TDI - Master IAD D - UPMC P6

Antoine MANZANERA - ENSTA/LEI

15

## Scale-space linéaire

Dans le cadre du scale-space linéaire, on identifie le *schéma de simplification* au filtrage passe-bas, et le *paramètre d'échelle* indique l'*amplitude spatiale* du filtre :

scale-space :  $(I_t)_{t \geq 0}$   
 avec :  $I_t = f_t(I_0)$  (linéaire)  
 $= I_0 * h_t$  (convolution)  
 donc  
 $f_t(I_0)(z) = \int_{x \in \mathbb{R}^2} I_0(z-x) h_t(x) dx$   
 où  $h_t$  est une fonction filtre passe-bas, d'amplitude spatiale  $t$

Axiomatique:

- Invariance en translation
- Invariance en rotation
- Invariance en échelle
- Notion de semi-groupe :

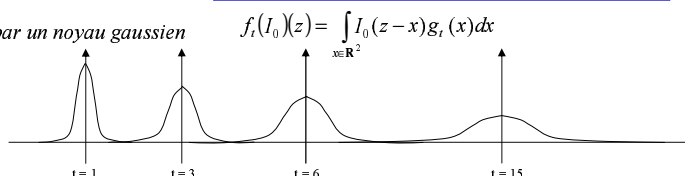
$$f_{t_1} \circ f_{t_2}(I) = f_{t_1+t_2}(I)$$

soit  $f_{t_2}(I) = f_{t_2-t_1}(f_{t_1}(I))$

$f_t$  = convolution par un noyau gaussien

avec  $g_t(x) = \frac{1}{\pi} e^{-x^2/t}$

$t$  : variance de la gaussienne



Cours TDI - Master IAD D - UPMC P6

Antoine MANZANERA - ENSTA/LEI

16



## Scale-space linéaire et noyau gaussien

Linéarité + Invariance en translation



Convolution :  
 $I(x, t) = I(x, 0) * h(x, t)$

Dans le domaine de Fourier :  $\hat{I}(u, t) = \hat{I}(u, 0) \cdot \hat{h}(u, t)$

Invariance en rotation



Fréquence scalaire :  
 $u = \|u\|$

Invariance en échelle



Equation aux dimensions :

$$\left[ \frac{\hat{I}(u, t)}{\hat{I}(u, 0)} \right] = [u \cdot t]$$



$$\frac{\hat{I}(u, t)}{\hat{I}(u, 0)} = \hat{h}(|u|t)^p$$

Relation de semi-groupe



$$\hat{h}(|u_1|^p) \hat{h}(|u_2|^p) = \hat{h}(|u_1|^p + |u_2|^p)$$



$$\hat{h}(u, t) = e^{-|u|^p t / 2}$$

Séparabilité



$$\hat{h}(u, t) = \exp\left(-\frac{1}{2} u^2 t^2\right)$$



Dans le domaine spatial :

$$h(x, t) = \frac{1}{2\pi^2} \exp\left(-\frac{1}{2t^2} x \cdot x\right)$$

[ter Haar Romeny 1994]

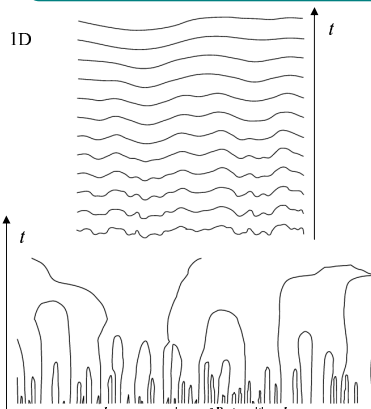
Cours TDI - Master IA&D - UPMC P6

Antoine MANZANERA - ENSTA/LEI

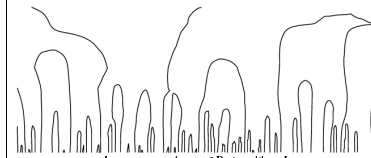
17

## Scale-space gaussien

1D



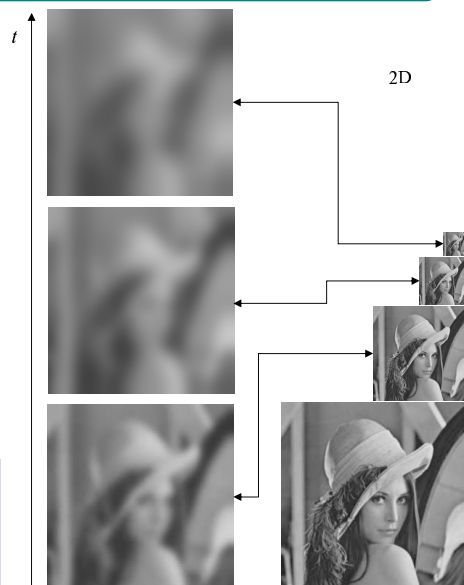
t



scale-space gaussien en 1D et position des extrema dans l'espace des échelles

Axiomatique du principe de causalité :

- Non création de nouveaux extrema locaux
- Non amplification des extrema existant



2D

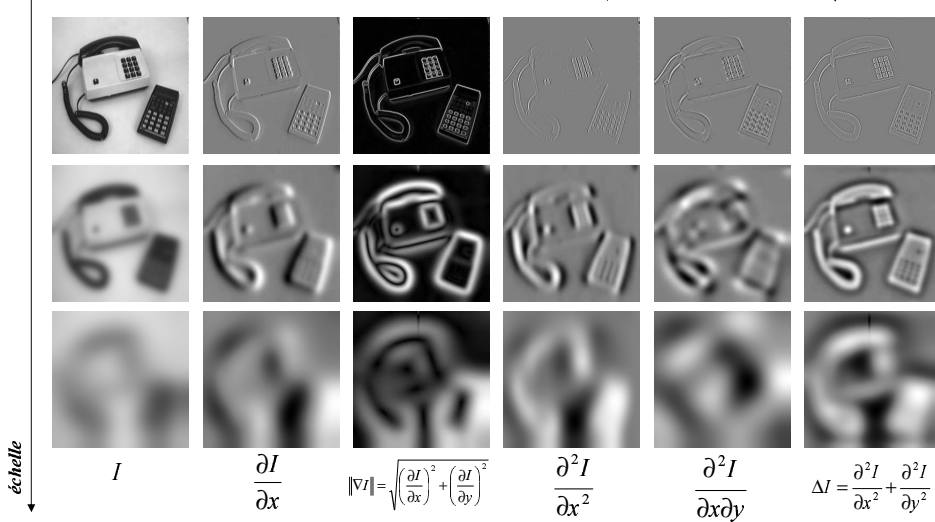
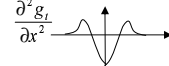
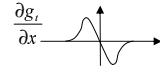
Cours TDI - Master IA&D - UPMC P6

Antoine MANZANERA - ENSTA/LEI

18

## Scale-space gaussien et dérivées

$$\partial^n (I * g_t) = I * \partial^n g_t$$



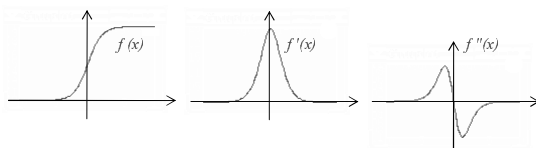
Cours TDI – Master I&A D – UPMC P6

Antoine MANZANERA – ENSTA/LEI

19

## Exemple 1 : Détection de contours

En 1d, un contour correspond à un maximum de la dérivée première, ou à un passage par zéro de la dérivée seconde :



En 2d, les contours se calculent à partir des dérivées premières (module et direction du gradient,...), ou des dérivées secondes (laplacien,...) :

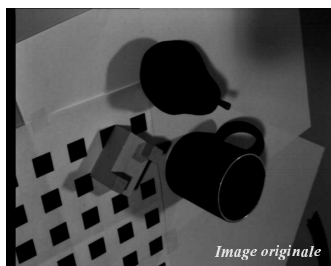
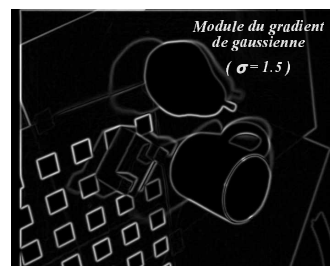
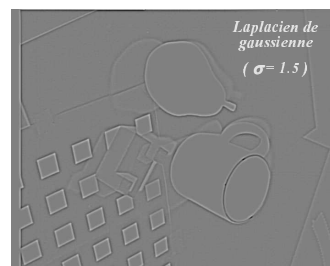


Image originale



Module du gradient de gaussienne (σ = 1.5)



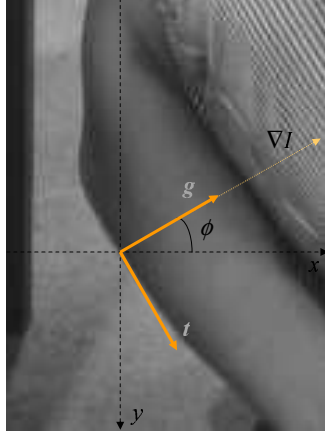
Laplacien de gaussienne (σ = 1.5)

Cours TDI – Master I&A D – UPMC P6

Antoine MANZANERA – ENSTA/LEI

20

## Exemple 1 : Détection de contours



On peut définir les *contours* dans une image  $I$  par le lieu des *maxima du module du gradient*  $\|\nabla I\|$ , dans la direction du gradient  $\arg(\nabla I)$ .

En chaque pixel  $(i, j)$  on peut définir un *système de coordonnées local*  $(g, t)$  à partir du gradient  $\nabla I(i, j)$ , comme suit :

$$\begin{cases} x = g \cos \phi + t \sin \phi \\ y = g \sin \phi - t \cos \phi \end{cases} \quad \begin{cases} g = x \cos \phi + y \sin \phi \\ t = -x \sin \phi + y \cos \phi \end{cases}$$

avec  $\phi = \arg(\nabla I) = \arctg(I_y/I_x)$  en notant :  $I_x = \frac{\partial I}{\partial x}$ ,  $I_{xx} = \frac{\partial^2 I}{\partial x^2}$ ,  $I_{xy} = \frac{\partial^2 I}{\partial x \partial y}$  ... etc

Les contours correspondent alors au passage par zéro de la dérivée seconde calculée dans la direction du gradient :

soit :  $I_{gg} = 0$

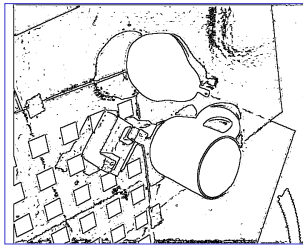
$$\text{or : } \begin{cases} I_g = I_x \cos \phi + I_y \sin \phi \\ I_t = I_x \sin \phi - I_y \cos \phi \end{cases} \quad \text{et : } \begin{cases} I_{gg} = I_{xx} \cos^2 \phi + 2I_{xy} \cos \phi \sin \phi + I_{yy} \sin^2 \phi \\ I_{tt} = I_{xx} \sin^2 \phi - 2I_{xy} \cos \phi \sin \phi + I_{yy} \cos^2 \phi \end{cases}$$

D'autre part, la dérivée seconde calculée dans la direction orthogonale au gradient correspond à la *courbure*  $I_{tt}$ . De plus :

$$\Delta I = I_{xx} + I_{yy} = I_{gg} + I_{tt} \quad (\text{le laplacien est invariant en rotation})$$

Donc, lorsque la courbure est faible, on peut assimiler le lieu des maxima du gradient dans la direction du gradient au lieu des passages par zéro du laplacien.

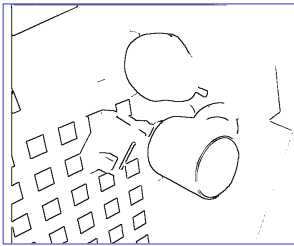
## Exemple 1 : Détection de contours



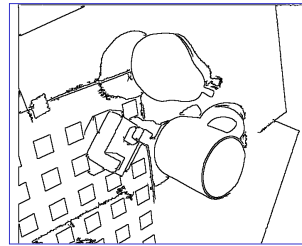
Passage par zéro du laplacien avec seuillage du module du gradient ( $s=1$ ) :

$$C_s(I) = \delta_B(\Delta_s I) \cap \delta_B(\Delta_{-s} I) \cap \sigma_s(\|\nabla I\|)$$

$$\text{avec : } \begin{cases} \Delta_s I = \{(x, y) / \Delta I(x, y) > 0\} \\ \Delta_{-s} I = \{(x, y) / \Delta I(x, y) < 0\} \\ \delta_B(X) = \{z - b / z \in X; b \in B\} \\ \quad \text{(dilatation morphologique par la} \\ \quad \text{boule élémentaire } B) \\ \sigma_s(X) = \{(x, y) / X(x, y) > s\} \end{cases}$$



Idem, avec ( $s=8$ )



Idem, avec seuillage par hystérésis du module du gradient ( $s_1=1, s_2=8$ ) :

$$C_{s_1}^{s_2}(I) = R_{C_{s_1}(I)}(C_{s_2}(I))$$

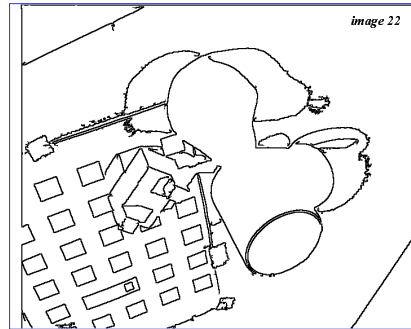
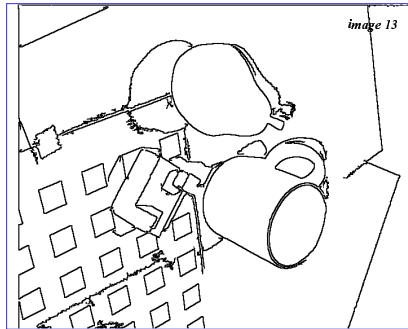
Reconstruction géodésique du seuillage haut dans le seuillage bas.

[Marr – Hildreth 80]

## Détection de contours : applications

La mise en correspondance de structures extraites par le détecteur de contours dans différentes images peut fournir des renseignements de haut niveau sur la scène :

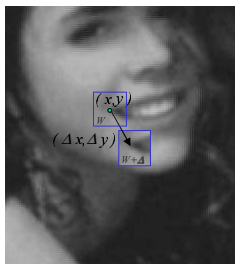
- |            |  |                |   |
|------------|--|----------------|---|
| exemples : | <ul style="list-style-type: none"> <li>• Déplacement de la caméra</li> <li>• Orientation des facettes</li> <li>• Profondeur</li> </ul> | Applications : | <ul style="list-style-type: none"> <li>• Navigation / guidage de robots.</li> <li>• Indexation d'images.</li> </ul> |
|------------|--|----------------|---|



Détection de passage par zéro du LoG ( $\sigma=1,5$ ) avec seuillage par hystérésis du module du gradient de gaussienne ( $s_1=1, s_2=8$ ).

## Exemple 2 : Détection de points anguleux

Les points anguleux (ou points d'intérêt, points saillants,...) sont des points « qui contiennent beaucoup d'information » relativement à l'image. Ce sont des points aux voisinages desquels l'image varie significativement dans plusieurs directions.



Une mesure des variations locales de l'image  $I$  au point  $(x, y)$  associée à un déplacement  $(\Delta x, \Delta y)$  est fournie par la fonction d'autocorrélation :

$$\chi(x, y) = \sum_{(x_k, y_k) \in W} (I(x_k, y_k) - I(x_k + \Delta x, y_k + \Delta y))^2$$

Où  $W$  est une fenêtre centrée au point  $(x, y)$ .

Or, en utilisant une approximation du premier ordre :

$$I(x_k + \Delta x, y_k + \Delta y) \approx I(x_k, y_k) + \frac{\partial I}{\partial x}(x_k, y_k) \cdot \frac{\Delta x}{\Delta y} + \frac{\partial I}{\partial y}(x_k, y_k) \cdot \frac{\Delta y}{\Delta x}$$

Et donc :

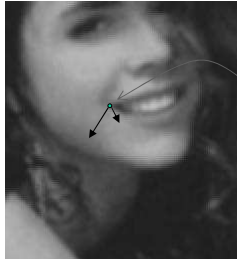
$$\chi(x, y) = \sum_{(x_k, y_k) \in W} \left( \frac{\partial I}{\partial x}(x_k, y_k) \cdot \frac{\Delta x}{\Delta y} + \frac{\partial I}{\partial y}(x_k, y_k) \cdot \frac{\Delta y}{\Delta x} \right)^2 = (\Delta x \quad \Delta y) \underbrace{\begin{pmatrix} \sum_{(x_k, y_k) \in W} \left( \frac{\partial I}{\partial x}(x_k, y_k) \right)^2 & \sum_{(x_k, y_k) \in W} \frac{\partial I}{\partial x}(x_k, y_k) \cdot \frac{\partial I}{\partial y}(x_k, y_k) \\ \sum_{(x_k, y_k) \in W} \frac{\partial I}{\partial x}(x_k, y_k) \cdot \frac{\partial I}{\partial y}(x_k, y_k) & \sum_{(x_k, y_k) \in W} \left( \frac{\partial I}{\partial y}(x_k, y_k) \right)^2 \end{pmatrix}}_{\Xi(x, y)} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix}$$

Matrice d'autocorrélation de l'image  $I$  en  $(x, y)$

## Exemple 2 : Détection de points anguleux

$$\Xi(x, y) = \begin{pmatrix} \sum_{(x_k, y_k) \in W} \left( \frac{\partial I}{\partial x}(x_k, y_k) \right)^2 & \sum_{(x_k, y_k) \in W} \frac{\partial I}{\partial x}(x_k, y_k) \cdot \frac{\partial I}{\partial y}(x_k, y_k) \\ \sum_{(x_k, y_k) \in W} \frac{\partial I}{\partial x}(x_k, y_k) \cdot \frac{\partial I}{\partial y}(x_k, y_k) & \sum_{(x_k, y_k) \in W} \left( \frac{\partial I}{\partial y}(x_k, y_k) \right)^2 \end{pmatrix}$$

La matrice d'autocorrélation  $\Xi$  représente la variation locale de l'image  $I$  en  $(x, y)$ .  $(x, y)$  sera considéré comme un point anguleux de  $I$  si pour tous les déplacements  $(\Delta x, \Delta y)$ , la quantité  $(\Delta x, \Delta y) \cdot \Xi(x, y) \cdot (\Delta x, \Delta y)^t$  est grande.



Les points anguleux sont les points  $(x, y)$  pour lesquels la matrice d'autocorrélation  $\Xi(x, y)$  a *deux valeurs propres grandes*.

Cela correspond aux points pour lesquels il existe localement une base de vecteurs propres de  $\Xi$  décrivant des variations locales importantes de l'image.

Le détecteur de Harris calcule une fonction d'intérêt  $\Theta(x, y)$  :

$$\Theta(x, y) = \det \Xi - \alpha \text{trace} \Xi$$

Le premier terme correspond au produit des valeurs propres, le second terme pénalise les points de contours avec une seule forte valeur propre.

Les points d'intérêt correspondent aux maxima locaux de la fonction  $\Theta$  qui sont au delà d'un certain seuil (typiquement 1% de la valeur max de  $\Theta$ ).

[Harris 88]

## Mise en oeuvre du détecteur de Harris

1. On calcule les dérivées premières à partir des dérivées de gaussienne (écart-type  $\sigma_1$ )
2. On calcule les termes de la matrice d'autocorrélation  $\Xi$  en calculant une moyenne locale des dérivées sous la forme d'une gaussienne (écart-type  $\sigma_2$ , typiquement  $\sigma_2 = 2 \sigma_1$ )
3. On calcule la fonction d'intérêt :  $\Theta = \det(\Xi) - \alpha \text{trace}(\Xi)$  (typiquement  $\alpha = 0,06$ ).
4. On calcule les maxima locaux de  $\Theta$  supérieurs à un certain seuil (typiquement 1% de  $\Theta_{\max}$ ).



$\sigma_1$  est donc le paramètre d'échelle du détecteur de Harris, qui détermine la portée des opérations de dérivations et d'intégrations (moyennes).

## Points anguleux : applications

Calcul des maxima locaux seuillés de la fonction d'intérêt :



Applications : reconstruction, poursuite, indexation,...

Cours TDI - Master I&A D - UPMC P6

Antoine MANZANERA - ENSTA/LEI

27

## Illustration du principe de causalité



Passage par zéro du laplacien (contours)



Détection de points d'intérêt (ou points anguleux)

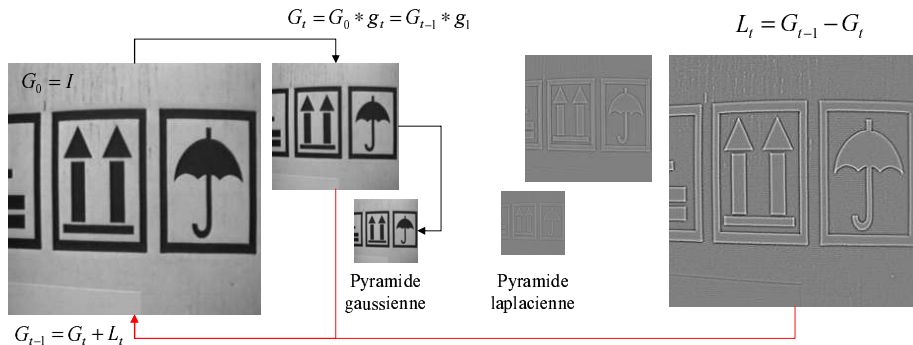
Cours TDI - Master I&A D - UPMC P6

Antoine MANZANERA - ENSTA/LEI

28

## Pyramides gaussienne et laplacienne

A une pyramide gaussienne on peut associer la « pyramide » des détails, c'est à dire la séquence des erreurs commises en approximant  $G_{t-1}$  par  $G_t$  :



La séquence des détails est appelée *pyramide laplacienne* (vient du fait que la différence de gaussiennes est une approximation du laplacien). La pyramide laplacienne fournit donc une *décomposition fréquentielle spatialement localisée* de l'image.

De plus, chaque niveau  $t$  de la pyramide laplacienne permet de reconstruire le niveau  $t-1$  de la pyramide gaussienne à partir de son niveau  $t$ .

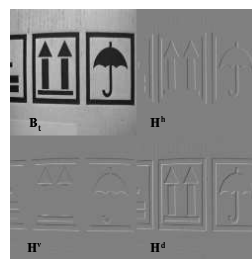
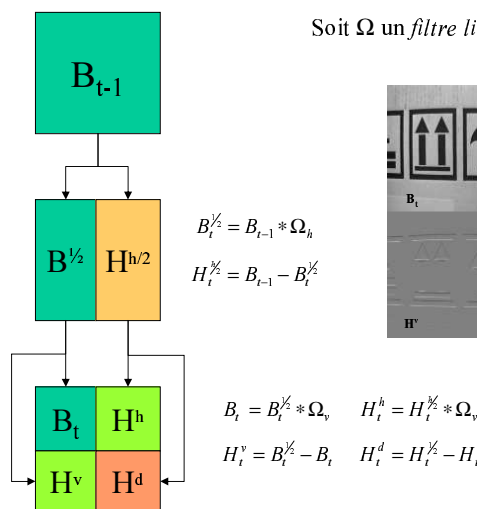
Cours TDI – Master IA&D – UPMC P6

Antoine MANZANERA – ENSTA/LEI

29

## Transformée en ondelettes

Soit  $\Omega$  un *filtre lisse*, séparable :  $\Omega = \Omega_h * \Omega_v$



La transformée en ondelettes fournit un codage de l'image en décomposant les valeurs de chaque pixel dans une base fournie par un couple de filtres (lisseur, dérivateur), dans deux directions orthogonales.

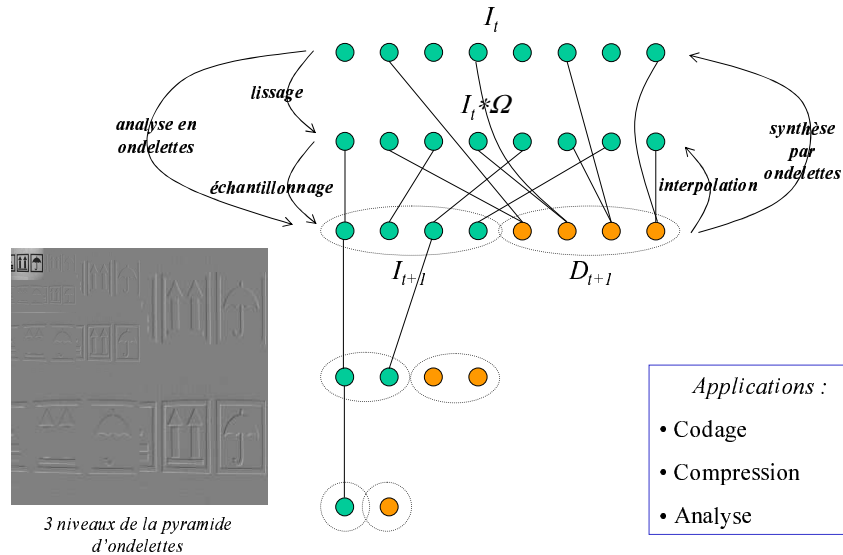
Attention, le principe de causalité n'est pas forcément respecté.

Cours TDI – Master IA&D – UPMC P6

Antoine MANZANERA – ENSTA/LEI

30

## Transformée en ondelettes



Applications :

- Codage
- Compression
- Analyse

## Application : compression d'images fixes

Comparaison entre la compression par transformée en cosinus discret (Jpeg) et la compression par transformée en ondelettes (Jpeg 2000); illustration du multi-échelles par la transmission progressive d'information image.



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 :**

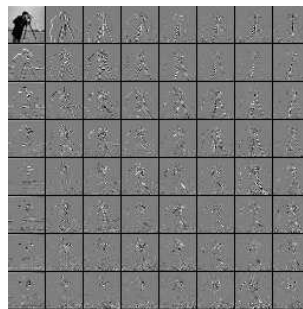
$$I(x, y) = \sum_{u=0}^{n-1} \sum_{v=0}^{n-1} \left( \frac{1}{\sqrt{2}} \right)^{\delta_{(0,0)}(u,v)} C(u, v) \cdot \cos \pi \frac{(2x+1)u}{n} \cdot \sin \pi \frac{(2y+1)v}{n}$$

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

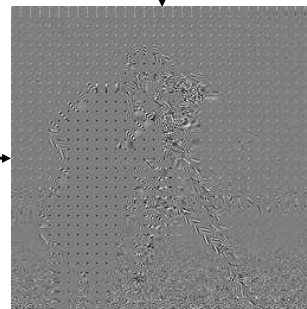
Le mode de représentation (1) de la DCT2d préserve la localisation des blocs de l'image originale : il illustre bien l'irrégularité spatiale de la répartition de l'information.

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.

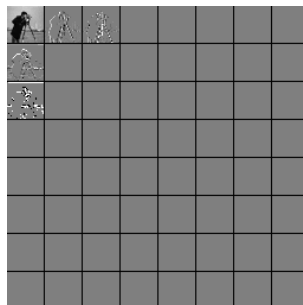
La norme Jpeg utilise en outre un codage entropique qui n'est pas pris en compte ci-dessous :

*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.

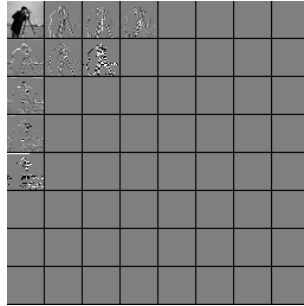
La norme Jpeg utilise en outre un codage entropique qui n'est pas pris en compte ci-dessous :

*originale*



256x256 pixels codés  
sur 8 bits.

*DCT2d quantifiée*



*DCT2d inverse*



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

Cours TDI – Master IA&D – UPMC P6

Antoine MANZANERA – ENSTA/LEI

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

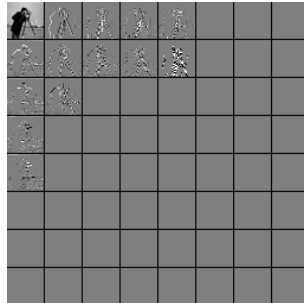
La norme Jpeg utilise en outre un codage entropique qui n'est pas pris en compte ci-dessous :

*originale*



256x256 pixels codés  
sur 8 bits.

*DCT2d quantifiée*



*DCT2d inverse*



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

Cours TDI – Master IA&D – UPMC P6

Antoine MANZANERA – ENSTA/LEI

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

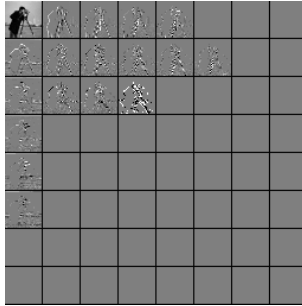
La norme Jpeg utilise en outre un codage entropique qui n'est pas pris en compte ci-dessous :

*originale*



256x256 pixels codés  
sur 8 bits.

*DCT2d quantifiée*



*DCT2d inverse*



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

Cours TDI – Master IA&D – UPMC P6

Antoine MANZANERA – ENSTA/LEI

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

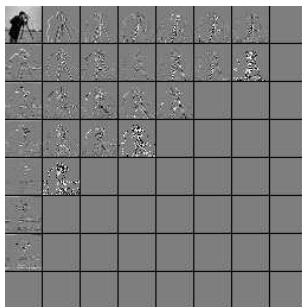
La norme Jpeg utilise en outre un codage entropique qui n'est pas pris en compte ci-dessous :

*originale*



256x256 pixels codés  
sur 8 bits.

*DCT2d quantifiée*



*DCT2d inverse*



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

Cours TDI – Master IA&D – UPMC P6

Antoine MANZANERA – ENSTA/LEI

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

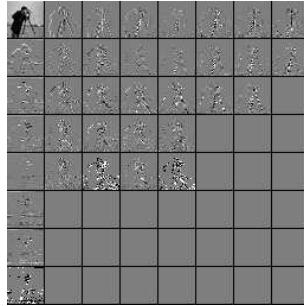
La norme Jpeg utilise en outre un codage entropique qui n'est pas pris en compte ci-dessous :

*originale*



256x256 pixels codés  
sur 8 bits.

*DCT2d quantifiée*



*DCT2d inverse*



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

Cours TDI – Master IA&D – UPMC P6

Antoine MANZANERA – ENSTA/LEI

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

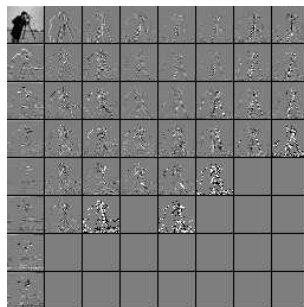
La norme Jpeg utilise en outre un codage entropique qui n'est pas pris en compte ci-dessous :

*originale*



256x256 pixels codés  
sur 8 bits.

*DCT2d quantifiée*



*DCT2d inverse*



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

Cours TDI – Master IA&D – UPMC P6

Antoine MANZANERA – ENSTA/LEI

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

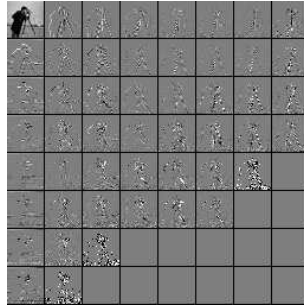
La norme Jpeg utilise en outre un codage entropique qui n'est pas pris en compte ci-dessous :

*originale*



256x256 pixels codés  
sur 8 bits.

*DCT2d quantifiée*



*DCT2d inverse*



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

Cours TDI – Master IA&D – UPMC P6

Antoine MANZANERA – ENSTA/LEI

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

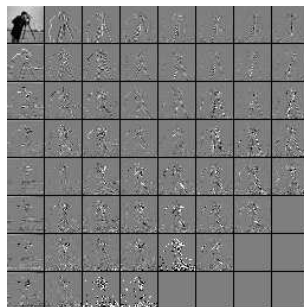
La norme Jpeg utilise en outre un codage entropique qui n'est pas pris en compte ci-dessous :

*originale*



256x256 pixels codés  
sur 8 bits.

*DCT2d quantifiée*



*DCT2d inverse*



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

Cours TDI – Master IA&D – UPMC P6

Antoine MANZANERA – ENSTA/LEI

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

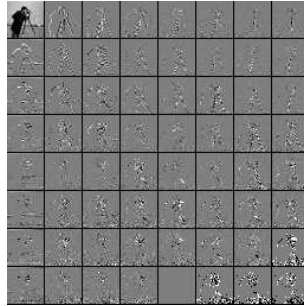
La norme Jpeg utilise en outre un codage entropique qui n'est pas pris en compte ci-dessous :

*originale*



256x256 pixels codés  
sur 8 bits.

*DCT2d quantifiée*



*DCT2d inverse*

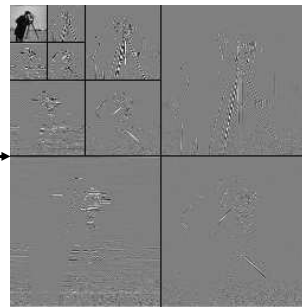


Bitrate : 4  
Size : 256 Ko  
Mean Square Error : 7,5

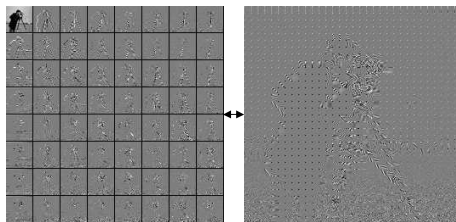
Cours TDI – Master IA&D – UPMC P6

Antoine MANZANERA – ENSTA/LEI

La norme de compression *Jpeg2000* est fondée sur l'algorithme EZW (Embedded Zero-tree Wavelets) qui opère une quantification des coefficients d'une *transformée en ondelettes* basée sur un schéma numérique (ondelettes de Daubechies).



*transformée en ondelettes (Daubechies)*



*DCT par blocs*

Par rapport à la DCT par blocs, la transformée en ondelettes est appliquée de façon uniforme sur toute l'image, ce qui évite les effets de blocs lors de la quantification des coefficients.

Le gain dans la compression vient principalement du fait que dans la DCT, la même quantification est appliquée pour tous les blocs quel que soit leur contenu. Cependant, cette différence est en partie compensée par le codage entropique. Au total, la taille d'une image Jpeg2000 par rapport à une image Jpeg de même qualité est entre 20% et 50% inférieure.

Cours TDI – Master IA&D – UPMC P6

Antoine MANZANERA – ENSTA/LEI

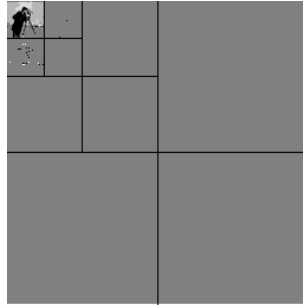
Transmission progressive d'image correspondant à une allocation de débit variable du codage en ondelettes opéré par l'algorithme EZW :

*originale*



256x256 pixels codés  
sur 8 bits.

*Transformée en  
ondelettes quantifiée*



*Transformée inverse*



Bitrate : 0,05  
Size : 3,2 Ko  
Mean Square Error : 752,9

Cours TDI – Master IA&D – UPMC P6

Antoine MANZANERA – ENSTA/LEI

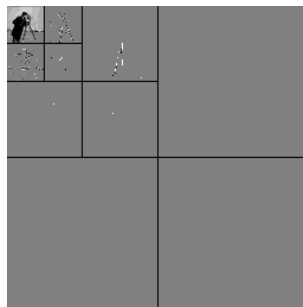
Transmission progressive d'image correspondant à une allocation de débit variable du codage en ondelettes opéré par l'algorithme EZW :

*originale*



256x256 pixels codés  
sur 8 bits.

*Transformée en  
ondelettes quantifiée*



*Transformée inverse*



Bitrate : 0,1  
Size : 6,4 Ko  
Mean Square Error : 534,0

Cours TDI – Master IA&D – UPMC P6

Antoine MANZANERA – ENSTA/LEI

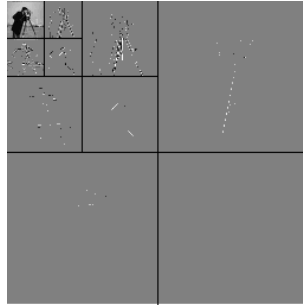
Transmission progressive d'image correspondant à une allocation de débit variable du codage en ondelettes opéré par l'algorithme EZW :

*originale*



256x256 pixels codés  
sur 8 bits.

*Transformée en  
ondelettes quantifiée*



*Transformée inverse*



Bitrate : 0,2  
Size : 12,8 Ko  
Mean Square Error : 249,6

Cours TDI – Master IA&D – UPMC P6

Antoine MANZANERA – ENSTA/LEI

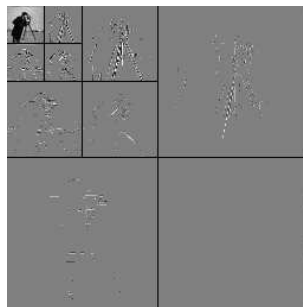
Transmission progressive d'image correspondant à une allocation de débit variable du codage en ondelettes opéré par l'algorithme EZW :

*originale*



256x256 pixels codés  
sur 8 bits.

*Transformée en  
ondelettes quantifiée*



*Transformée inverse*



Bitrate : 0,3  
Size : 19,2 Ko  
Mean Square Error : 141,2

Cours TDI – Master IA&D – UPMC P6

Antoine MANZANERA – ENSTA/LEI



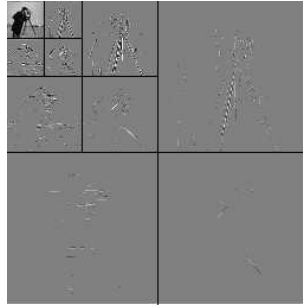
Transmission progressive d'image correspondant à une allocation de débit variable du codage en ondelettes opéré par l'algorithme EZW :

*originale*



256x256 pixels codés  
sur 8 bits.

*Transformée en  
ondelettes quantifiée*



*Transformée inverse*



Bitrate : 0,4  
Size : 25,6 Ko  
Mean Square Error : 113,1

Cours TDI – Master IA&D – UPMC P6

Antoine MANZANERA – ENSTA/LEI

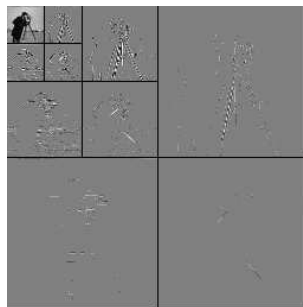
Transmission progressive d'image correspondant à une allocation de débit variable du codage en ondelettes opéré par l'algorithme EZW :

*originale*



256x256 pixels codés  
sur 8 bits.

*Transformée en  
ondelettes quantifiée*



*Transformée inverse*



Bitrate : 0,5  
Size : 32 Ko  
Mean Square Error : 89,4

Cours TDI – Master IA&D – UPMC P6

Antoine MANZANERA – ENSTA/LEI

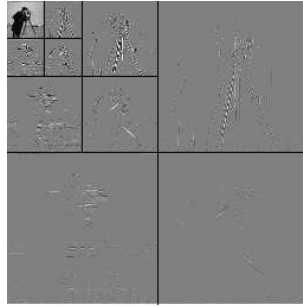
Transmission progressive d'image correspondant à une allocation de débit variable du codage en ondelettes opéré par l'algorithme EZW :

*originale*



256x256 pixels codés  
sur 8 bits.

*Transformée en  
ondelettes quantifiée*



*Transformée inverse*



Bitrate : 0,75  
Size : 48 Ko  
Mean Square Error : 47,7

Cours TDI – Master IA&D – UPMC P6

Antoine MANZANERA – ENSTA/LEI

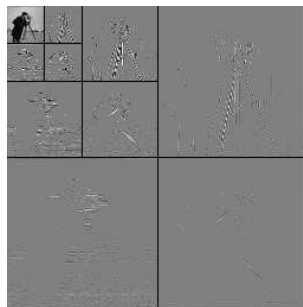
Transmission progressive d'image correspondant à une allocation de débit variable du codage en ondelettes opéré par l'algorithme EZW :

*originale*



256x256 pixels codés  
sur 8 bits.

*Transformée en  
ondelettes quantifiée*



*Transformée inverse*



Bitrate : 1  
Size : 64 Ko  
Mean Square Error : 26,4

Cours TDI – Master IA&D – UPMC P6

Antoine MANZANERA – ENSTA/LEI

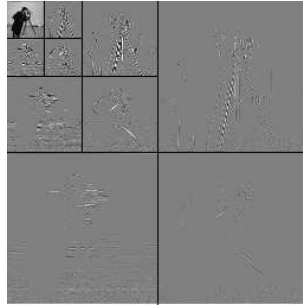
Transmission progressive d'image correspondant à une allocation de débit variable du codage en ondelettes opéré par l'algorithme EZW :

*originale*



256x256 pixels codés  
sur 8 bits.

*Transformée en  
ondelettes quantifiée*



*Transformée inverse*



Bitrate : 1,5  
Size : 96 Ko  
Mean Square Error : 11,7

Cours TDI – Master IA&D – UPMC P6

Antoine MANZANERA – ENSTA/LEI

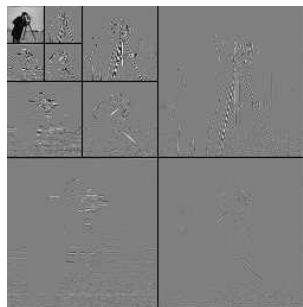
Transmission progressive d'image correspondant à une allocation de débit variable du codage en ondelettes opéré par l'algorithme EZW :

*originale*



256x256 pixels codés  
sur 8 bits.

*Transformée en  
ondelettes quantifiée*



*Transformée inverse*



Bitrate : 3  
Size : 192 Ko  
Mean Square Error : 1,8

Cours TDI – Master IA&D – UPMC P6

Antoine MANZANERA – ENSTA/LEI

## Scale-space gaussien et diffusion

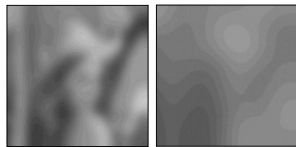
L'image convoluée avec une gaussienne de paramètre  $t$  est solution de l'équation de diffusion de la chaleur :

$$\text{solution : } \begin{cases} I = I_0 & \text{cond. initiale} \\ I_t = I_0 * g_t \end{cases}$$

$$\frac{\partial I}{\partial t} = \text{div}(c \nabla I) = c \Delta I$$

gradient temporel      divergence      conductance      gradient spatial      laplacien

Dans le cadre de l'équation de diffusion, il y a *identité* entre le *facteur de temps* et le *facteur d'échelle*.



diffusion de la chaleur



passages par zéro du laplacien

Dans ce cas le facteur de conductance  $c$  est constant sur toute l'image (diffusion isotrope).

Mais on peut rendre  $c$  variable, dépendant du contenu de l'image (diffusion anisotrope).

↓  
Espaces d'échelle non linéaires

## Diffusion anisotrope

Dans la diffusion isotrope, le facteur de conductance  $c$  est constant :

$$\frac{\partial I}{\partial t} = \text{div}(c \nabla I) = c \Delta I$$

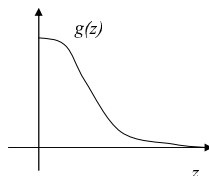
L'idée de la diffusion anisotrope, c'est de rendre  $c$  variable :

$$\frac{\partial I}{\partial t} = \text{div}(c \nabla I) = c \Delta I + \nabla c \cdot \nabla I$$

Et dépendant du contenu de l'image. Précisément, on souhaiterait que la diffusion soit favorisée dans les régions plutôt homogènes et réduite au voisinage des frontières de ces régions. On choisit typiquement une fonction décroissante du module du gradient :

$$c(x, y, t) = g(\|\nabla I(x, y, t)\|)$$

comportement souhaité de la fonction  $g$  :



Par exemple :  $c(x, y, t) = e^{-\left(\frac{\|\nabla I\|}{\kappa}\right)^2}$

ou bien :  $c(x, y, t) = \frac{1}{1 + \left(\frac{\|\nabla I\|}{\kappa}\right)^2}$

Contrairement au cas précédent, on ne connaît pas de solution explicite de l'équation d'évolution pour chaque échelle. Cette solution est donc calculée par un schéma numérique.

## Schéma numérique de diffusion anisotrope

Le schéma numérique proposé par les auteurs pour la discrétisation de l'équation de diffusion anisotrope est le suivant :

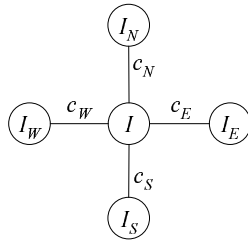
Le laplacien est discrétisé aux 4 plus proches voisins.

Les coefficients de conductance sont distribués sur les arcs de la maille carrée.

L'équation d'évolution discrète est la suivante :

$$I^{t+1} = I^t + \lambda [c_N \Delta_N I + c_E \Delta_E I + c_S \Delta_S I + c_W \Delta_W I]^t$$

avec :  $0 \leq \lambda \leq \frac{1}{4}$  (stabilité du schéma numérique)



avec :

$$\begin{cases} \Delta_N I = I_N - I \\ \Delta_E I = I_E - I \\ \Delta_S I = I_S - I \\ \Delta_W I = I_W - I \end{cases} \quad \text{et :} \quad \begin{cases} c_N = g(|\Delta_N I|) \\ c_E = g(|\Delta_E I|) \\ c_S = g(|\Delta_S I|) \\ c_W = g(|\Delta_W I|) \end{cases} \quad \begin{matrix} g(x) = e^{-\left(\frac{x}{K}\right)^2} \\ \text{ou} \\ g(x) = \frac{1}{1 + \left(\frac{x}{K}\right)^2} \end{matrix}$$

Rq<sub>1</sub> : lorsque  $c_N = c_E = c_S = c_W = 1$ , on retrouve  $I^{t+1} - I^t = \lambda (I_N + I_E + I_S + I_W - 4I)$ , schéma discret de diffusion isotrope.

Rq<sub>2</sub> : le paramètre  $K$  représente la valeur de gradient à partir de laquelle on limite la diffusion.

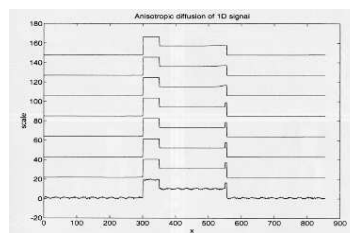
Valeur typique : 10% de la dynamique ou estimé automatiquement en fonction du bruit.

Lorsque  $K \rightarrow \infty$ , on tend vers un schéma isotrope.

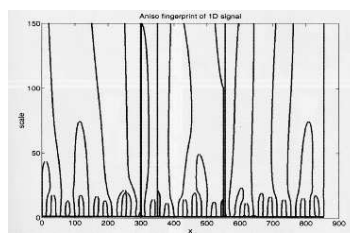
## La pyramide de diffusion anisotrope



$t = 0$



diffusion anisotrope d'un signal 1d



position des extrema dans l'espace d'échelles



$t = 10$



$t = 20$



$t = 60$



$t = 100$

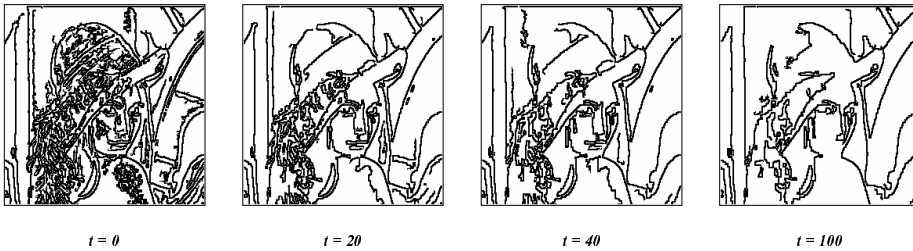
Diffusion anisotrope, schéma de décroissance hyperbolique (Image sur 8 bits,  $K=15$ ).

## Diffusion anisotrope et extrema

Diffusion anisotrope, schéma de décroissance hyperbolique (Image sur 8 bits,  $K=15$ ).



Passage par zéro du loG ( $\sigma=1,5$ ) avec seuil par hystérésis du module du gradient ( $S_{haut}=5; S_{bas}=1$ ).



Cours TDI – Master IA&D – UPMC P6

Antoine MANZANERA – ENSTA/LEI

59

## Espaces d'échelles morphologiques

Filtre morphologique : opérateur  $\Phi$  :

1. croissant :  $x \leq y \Rightarrow \Phi(x) \leq \Phi(y)$   $\longleftrightarrow$  simplification, perte d'information,...
2. idempotent :  $\Phi(\Phi(x)) = \Phi(x)$   $\longleftrightarrow$  stabilité, classes d'invariance,...

Dans les opérations de filtrage morphologique, la taille de l'élément structurant est en général un facteur d'échelle.

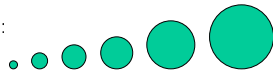
La notion de *semi-groupe* des espaces d'échelle se transpose dans les lois d'absorption liées aux granulométries.

granulométrie :

famille d'ouvertures  $(\gamma_\lambda)_{\lambda \geq 0}$  vérifiant :

$$0 \leq \lambda \leq \lambda' \Rightarrow \gamma_\lambda \gamma_{\lambda'} = \gamma_{\lambda'} \gamma_\lambda = \gamma_{\lambda'}$$

ex :



Ouvertures par des boules euclidiennes de rayon  $\lambda$

filtres alternés séquentiels :

$$\Theta_\lambda = \varphi_\lambda \gamma_\lambda \dots \varphi_2 \gamma_2 \varphi_1 \gamma_1$$

$$\Xi_\lambda = \gamma_\lambda \varphi_\lambda \dots \gamma_2 \varphi_2 \gamma_1 \varphi_1$$

$$\varphi_\lambda = (\gamma_\lambda)^*$$

Propriétés d'absorption :

$$\lambda \leq \lambda' \Rightarrow \begin{cases} \Theta_{\lambda'} \Theta_\lambda = \Theta_{\lambda'} \\ \Xi_{\lambda'} \Xi_\lambda = \Xi_{\lambda'} \end{cases}$$

Cours TDI – Master IA&D – UPMC P6

Antoine MANZANERA – ENSTA/LEI

60

## Premier scale-space morphologique

Une granulométrie induit un espace d'échelle morphologique via les filtres alternés séquentiels :



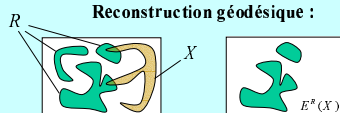
Cours TDI – Master IA&D – UPMC P6

Antoine MANZANERA – ENSTA/LEI

61

## Opérations géodésiques

**Reconstruction géodésique :**

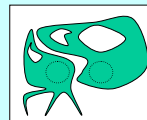


La reconstruction géodésique de  $R$  dans  $X$ , c'est l'ensemble des composantes connexes de  $R$  qui intersectent  $X$ .

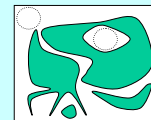
**L'ouverture par reconstruction** élimine les composantes connexes qui n'appartiennent pas à l'ouvert sans modifier les autres :



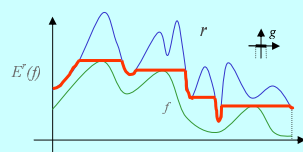
$X$



$E^X(Y_B(X))$



$(E^{X^c}((\phi_B(X))^c))^c$



**Reconstruction fonctionnelle :**

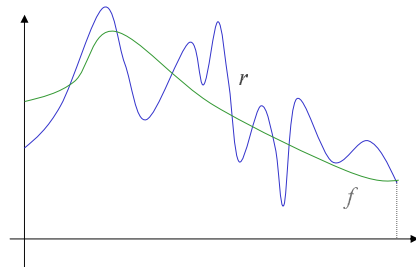
↓  
filtrage par *sélection* des ensembles de niveaux

Cours TDI – Master IA&D – UPMC P6

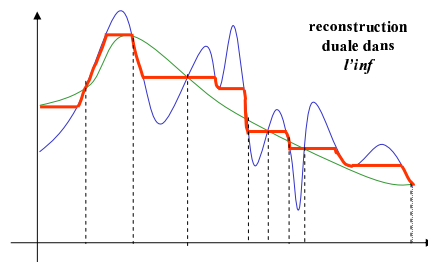
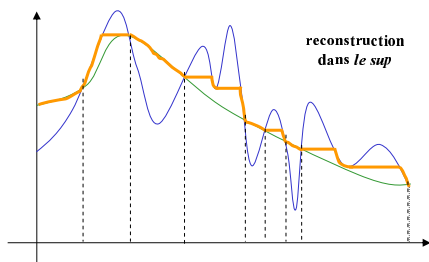
Antoine MANZANERA – ENSTA/LEI

62

## Nivellements



Les *nivellements* définissent des *opérateurs connexes*, qui simplifient l'image par *sélection* des ensembles de niveaux ou de leurs complémentaires :



Cours TDI – Master IA&D – UPMC P6

Antoine MANZANERA – ENSTA/LEI

63

## Second scale-space morphologique

Une granulométrie induit un espace d'échelle via les filtres alternés séquentiels par reconstruction :



Cours TDI – Master IA&D – UPMC P6

Antoine MANZANERA – ENSTA/LEI

64

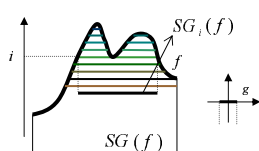


## Invariance par changement de contraste

$$f \leftrightarrow SG(f)$$

$$SG(f) = \bigcup_{i \in \mathbb{R}} SG_i(f) \times \{i\}$$

Par définition les opérateurs morphologiques croissants en niveaux de gris peuvent être obtenus à partir des opérateurs ensemblistes appliqués sur les sections du sous-graphe (ensembles de niveaux).



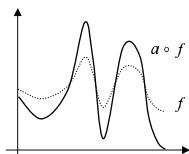
$$SG_i(f) = \{x \in \mathbb{R}^n / f(x) \geq i\}$$

$$\downarrow \text{érosion ensembliste}$$

$$SG_i(h) = \mathcal{E}_{\text{supp}(g)}(SG_i(f))$$

$$SG(h) = \bigcup_{i \in \mathbb{R}} SG_i(h) \times \{i\}$$

$$SG(h) \leftrightarrow h = \mathcal{E}_g(f)$$



Une conséquence de cette définition est l'invariance par changement de contraste : les opérateurs morphologiques commutent avec les anamorphoses, c'est-à-dire les transformations croissantes des niveaux de gris :

$$a \circ \Phi(f) = \Phi(a \circ f)$$

## Invariance par changement de contraste

L'invariance par changement de contraste est au cœur de l'axiomatique des espaces d'échelle morphologiques. Exprimé en termes d'équation aux dérivées partielles, on montre que dans un espace d'échelles invariant par contraste (i.e. un espace d'échelles morphologique), la diffusion se fait nécessairement dans la direction de la courbure des ensembles de niveaux :

On a une équation de la forme :

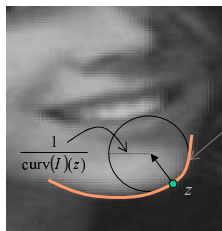
$$\frac{\partial I}{\partial t} = \|\nabla I\| G(\text{curv}(I), t)$$

où l'on note

$$\begin{cases} I_x = \frac{\partial I}{\partial x} & I_{xx} = \frac{\partial^2 I}{\partial x^2} & I_{xy} = \frac{\partial^2 I}{\partial x \partial y} \\ I_y = \frac{\partial I}{\partial y} & I_{yy} = \frac{\partial^2 I}{\partial y^2} \end{cases}$$

Avec :  $\text{curv}(I) = \text{div} \left( \frac{\nabla I}{\|\nabla I\|} \right) = \frac{I_{xx} I_y^2 - 2 I_{xy} I_x I_y + I_{yy} I_x^2}{(I_x^2 + I_y^2)^{3/2}}$

et avec  $G(x, y)$  continue et croissante par rapport à  $x$ .



La courbure de  $I$  au point  $z$  est égale à l'inverse du rayon du cercle osculateur à la courbe isophote en  $z$ , c'est-à-dire à la courbe de niveau  $I^{I(z)} = \{(x, y) / I(x, y) = I(z)\}$

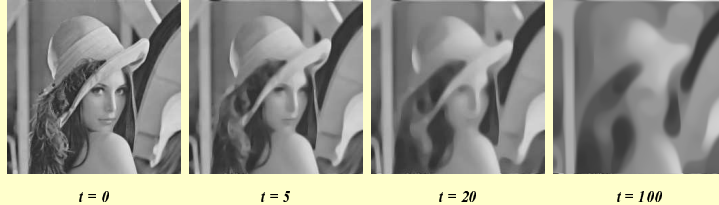
L'intérêt du formalisme EDP est de généraliser les filtres morphologiques et de fournir un cadre rigoureux aux transformations utilisant des éléments structuraux infinitésimaux.

[Alvarez et al 92]

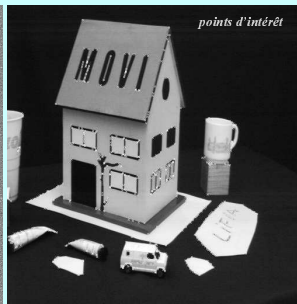
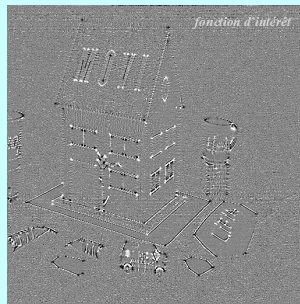
## Diffusion par la courbure

Une des équations de diffusion invariante par changement de contraste la plus simple est donc la diffusion par courbure moyenne :

$$\frac{\partial I}{\partial t} = \|\nabla I\| \text{curv}(I)$$



Application : points d'intérêt morphologiques



$$S(i) = \|\nabla I\| \text{curv}(I)$$

La fonction d'intérêt (ou de saillance) morphologique  $S(I)$  peut être définie comme le résidu de la diffusion par courbure moyenne ; elle est égale au produit du gradient par la courbure locale à l'isophote. Les points d'intérêts sont les maxima locaux supérieurs à un certain seuil de cette fonction d'intérêt.

Cours TDI – Master IA&D – UPMC P6

Antoine MANZANERA – ENSTA/LEI

67

## Squelettes multi-échelles

Le *squelette* est un *descripteur de formes* binaires lié à la notion d'*axe médian*, qui correspond à l'ensemble des *centres des boules maximales*.

On peut aussi définir le squelette comme le lieu des *chocs* lors de la *propagation* du contour de la forme binaire. L'*intensité* du choc se mesure par la différence des labels propagés, le label correspondant au numéro de l'élément de contour le plus proche, numéro fourni par une énumération cyclique du contour.

L'intensité du choc est un *facteur d'échelle* qui induit un espace d'échelle dans la description des formes binaires.

Costa et al., 1999

<http://cyvision.if.sc.usp.br/msskeletons/>



propagation des labels pour le calcul du squelette multi-échelles

Cours TDI – Master IA&D – UPMC P6

Antoine MANZANERA – ENSTA/LEI

68

## *Bibliographie du chapitre*

- **D. Marr & E. Hildreth 1980 « Theory of edge detection »** *Journal of Royal Society London B-207* 187-217
- **C. Harris & M. Stephens 1988 « A combined corner and edge detector »** *Alvey Vision Conference* pp 147-151
- **B.M. ter Haar Romeny, ed. 1994 « Geometry-driven diffusion in computer vision »** Kluwer Academic Publishers
- **J.M. Jolion & A. Rosenfeld 1994 « A pyramid framework for early vision »** Kluwer Academic Publishers
- **T. Lindeberg 1996 « Scale-space: a framework for handling image structures at multiple scales »** in *Proc. CERN School of computing*
- **J. Serra 1988 « Mathematical morphology and image analysis Vol. 2: theoretical advances »** Academic Press NY