

Traitement d'images et vision

Introduction aux espaces d'échelle

Cours D9-2 – ENSTA 3^e année

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*
 - Espaces d'échelles et diffusion
 - Espaces d'échelles non linéaires

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:



feuille



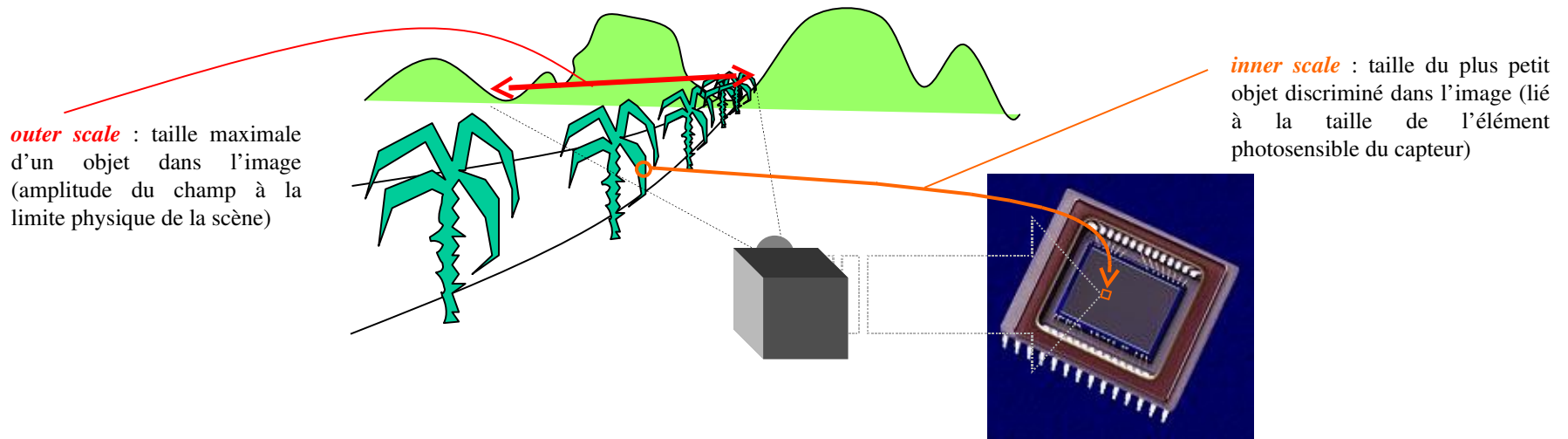
arbre



forêt

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]

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) :



$f[x, y]$



$\Delta f[x, y]$



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

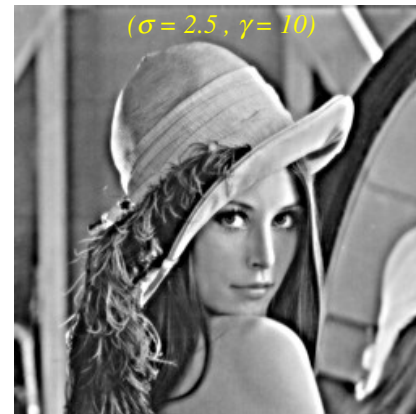
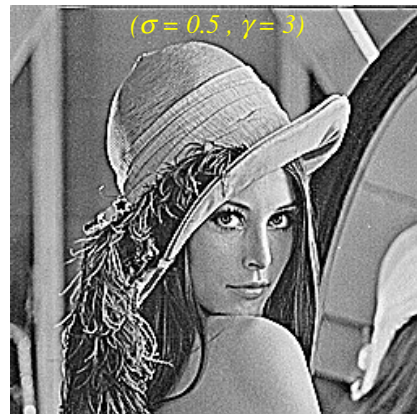
(γ : gain)

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

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

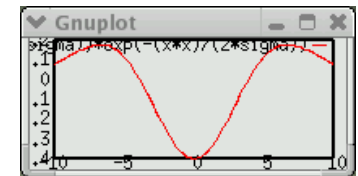
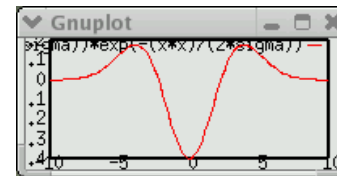
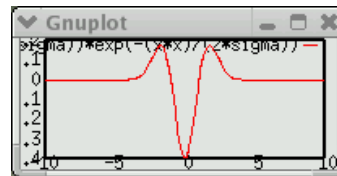
Introduction : Portée des opérateurs de TI

Mais le contraste est une notion multi-échelle :



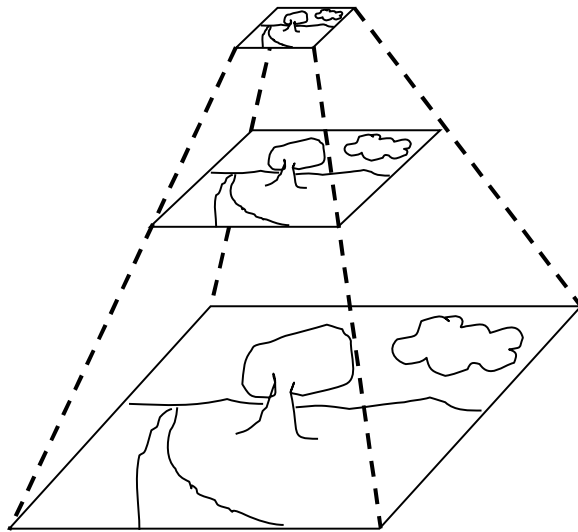
Il faut donc des mécanismes pour adapter la portée des opérateurs.

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* :



- Au niveau *abstrait* :

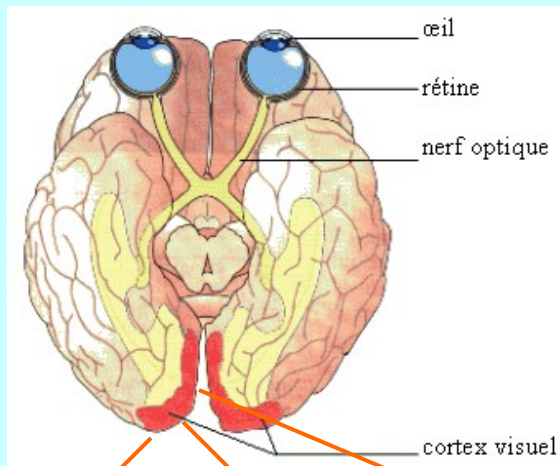


Intérêts algorithmiques du calcul multi-échelle :

- Robustesse
- Réduction du temps de calcul

Liens avec la vision biologique

« Codage » dans les canaux
fréquentiels corticaux



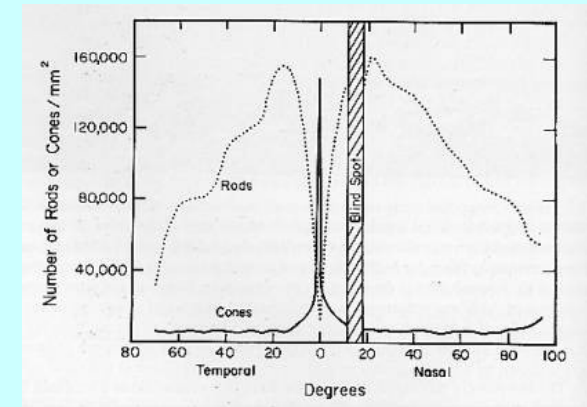
grossier

moyen

fin

Vision rétinienne : acuité décroissante du
centre à la périphérie

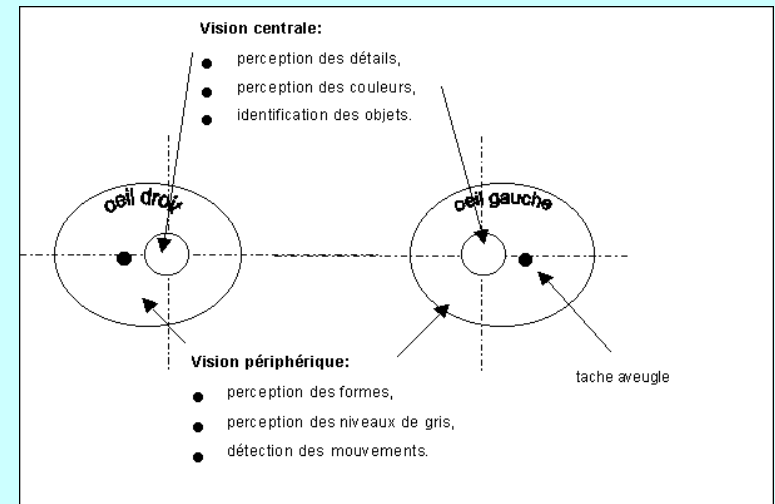
distribution inhomogène
des photorécepteurs de la
rétine (cônes et bâtonnets) :



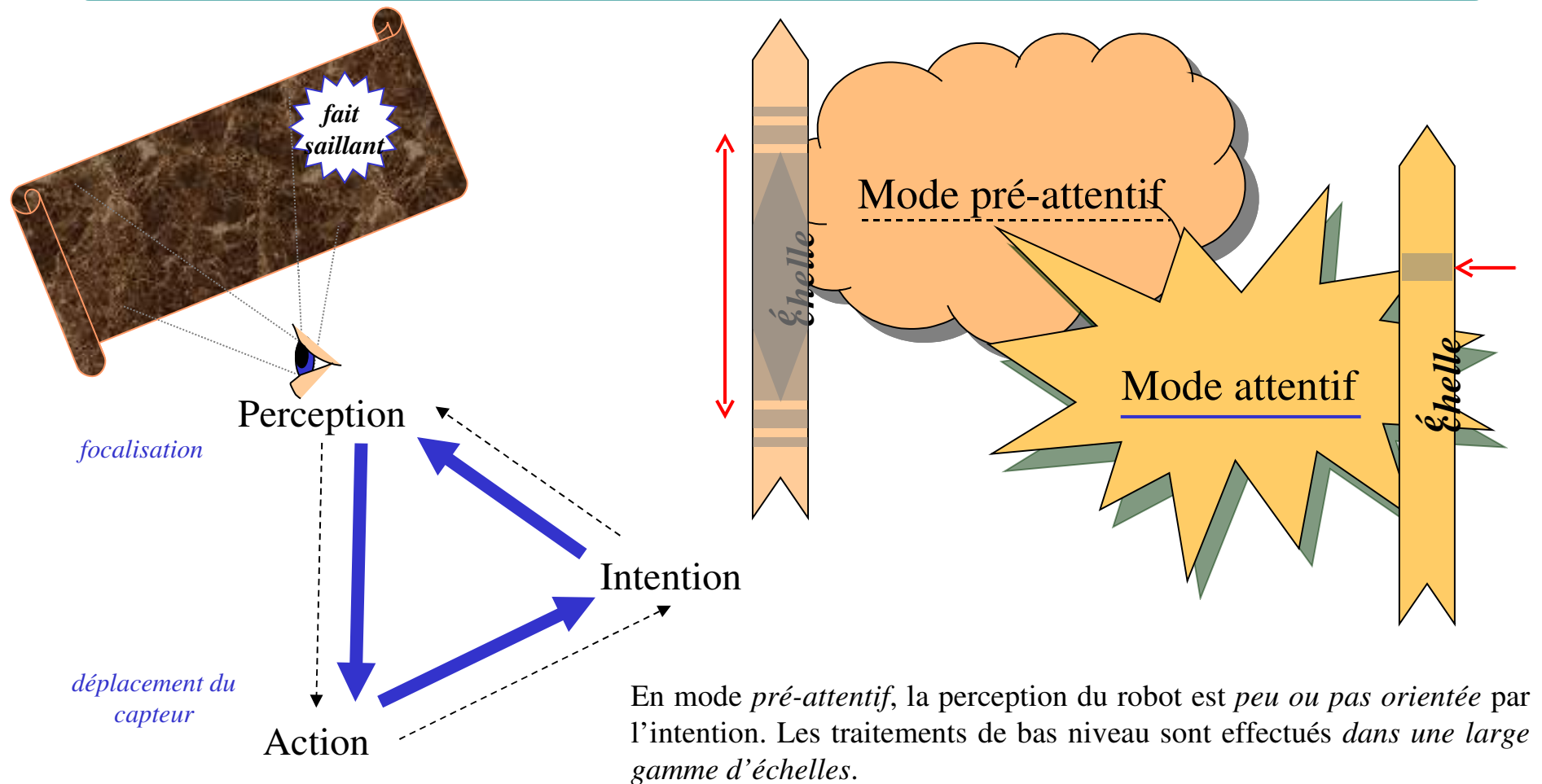
- focus d'attention
- zones préattentives



- détection du mouvement
- tracking
- vision active



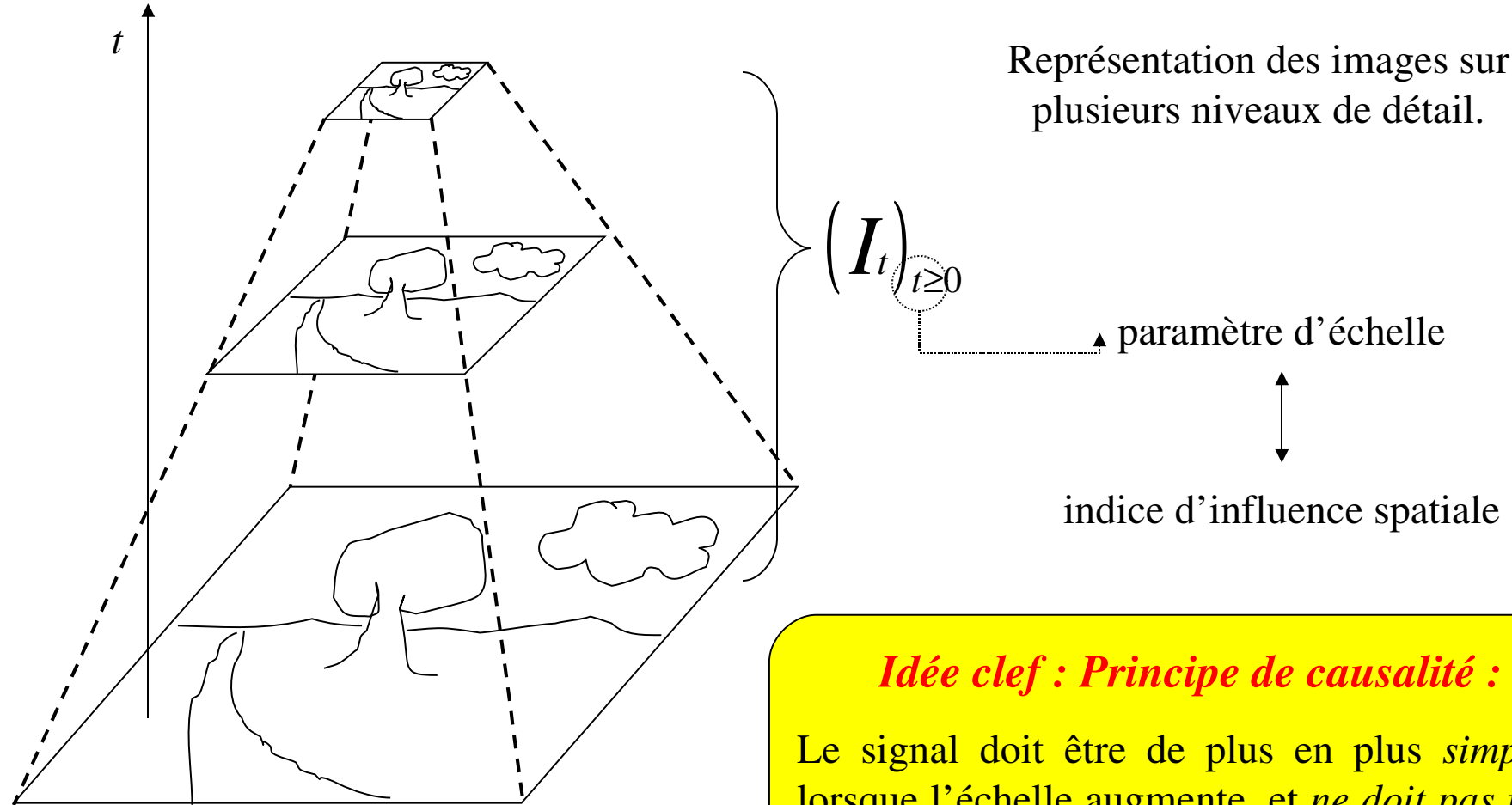
Espaces d'échelles et vision active



La pyramide visuelle

PYRAMIDE :

Représentation des images sur plusieurs niveaux de détail.



Idée clef : Principe de causalité :

Le signal doit être de plus en plus *simplifié* lorsque l'échelle augmente, et *ne doit pas faire apparaître de nouvelles structures*.

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.

Pyramides et sous-échantillonnage



Illustration en 2D :

Pyramide de Nyquist



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 :

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

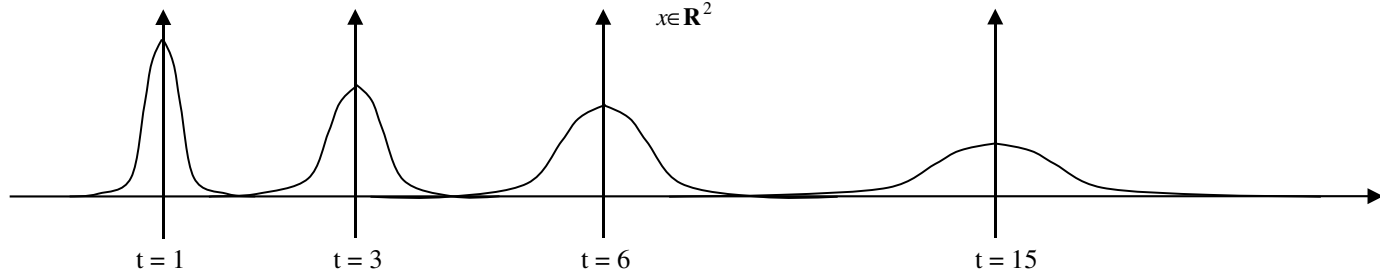
soit $f_{t_2}(I) = f_{t_2-t_1}(f_{t_1}(I))$

→ $f_t = \text{convolution par un noyau gaussien}$

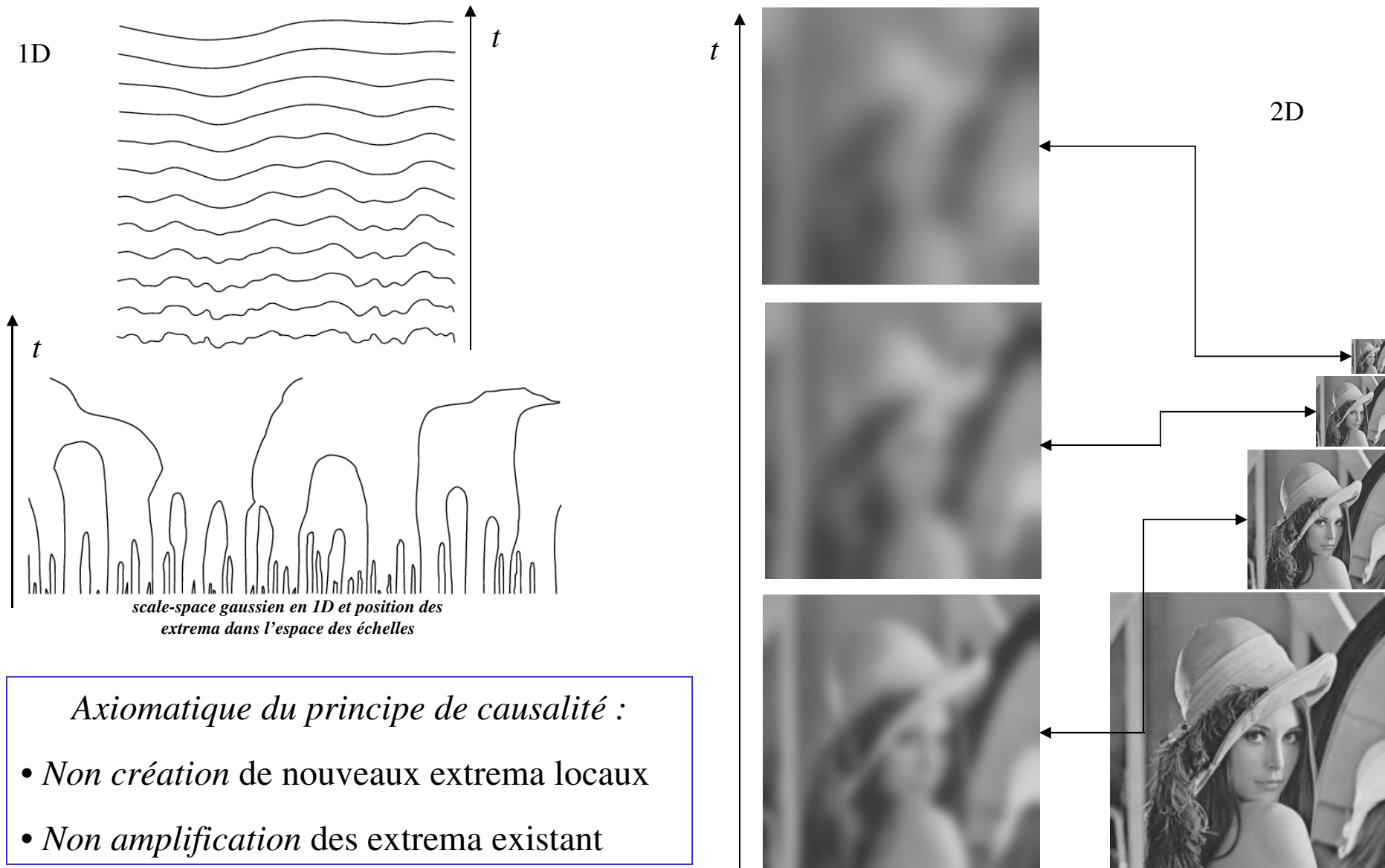
avec $g_t(x) = \frac{1}{\pi t} e^{-x^2/t}$

t : variance de la gaussienne

$$f_t(I_0)(z) = \int_{x \in \mathbb{R}^2} I_0(z-x) g_t(x) dx$$

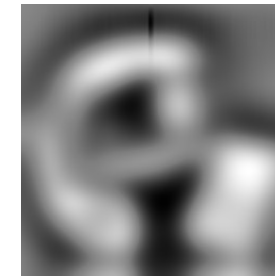
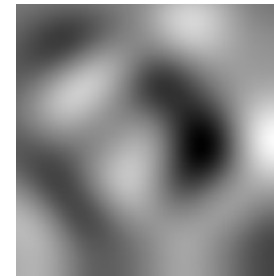
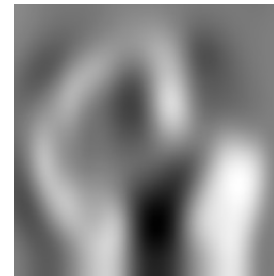
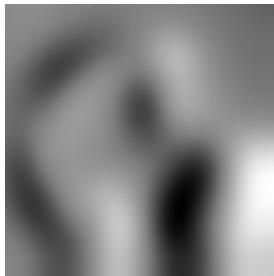
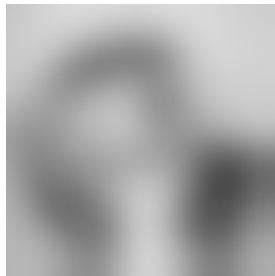
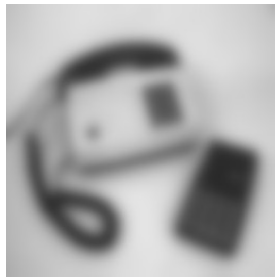
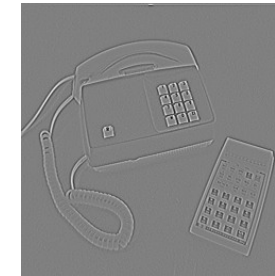
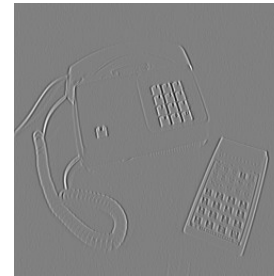
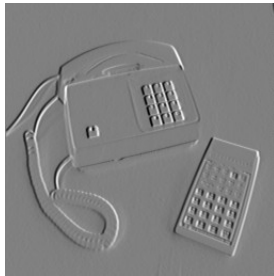
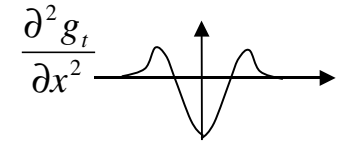
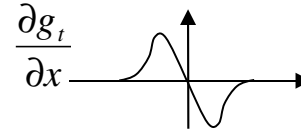
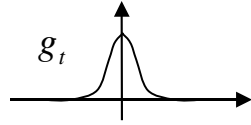


Scale-space gaussian



Scale-space gaussien et dérivées

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



I

$$\frac{\partial I}{\partial x}$$

$$\|\nabla I\| = \sqrt{\left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial y}\right)^2}$$

$$\frac{\partial^2 I}{\partial x^2}$$

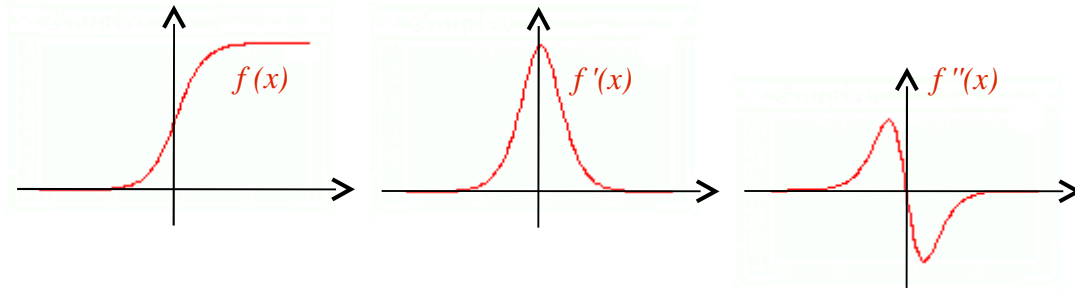
$$\frac{\partial^2 I}{\partial x \partial y}$$

$$\Delta I = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2}$$

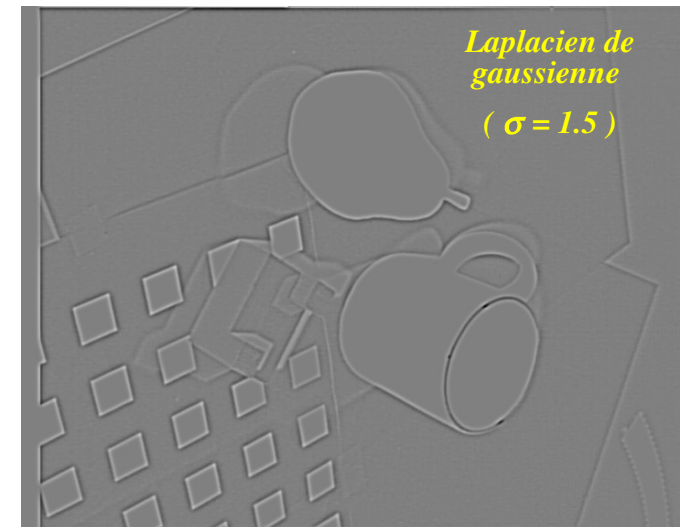
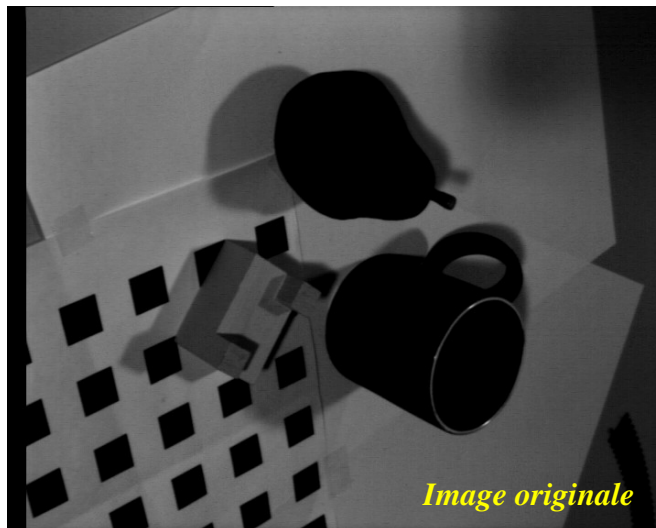
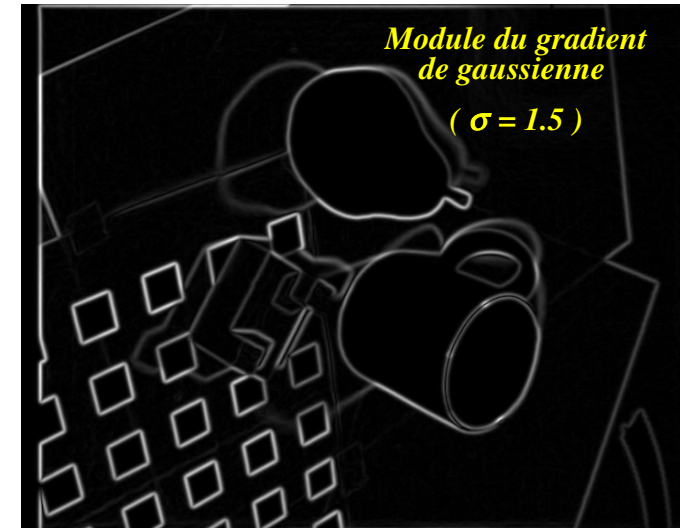
échelle
↓

Exemple 1 : Détection de contours

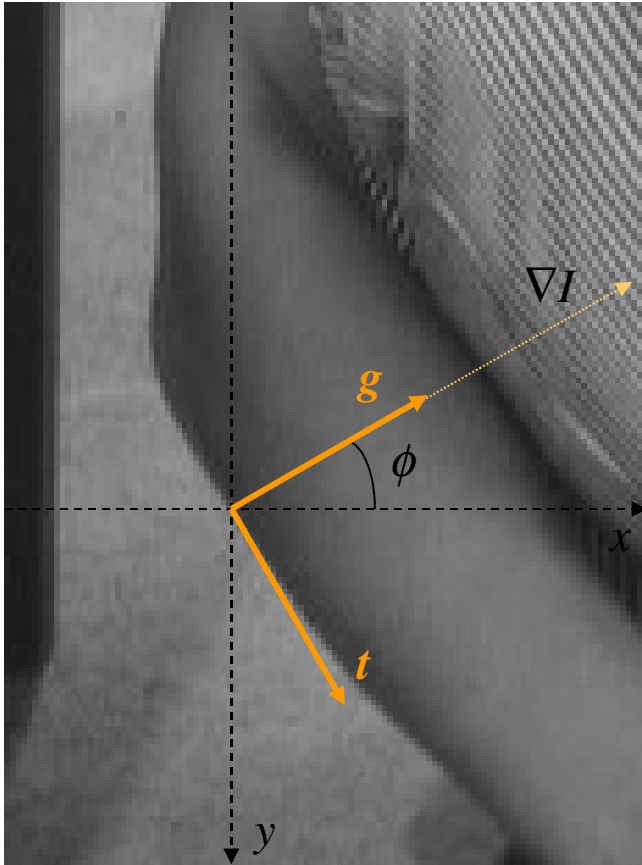
En 1d, un contour correspond à un maximum de la dérivée première, ou à un passage par zéro de la dérivée seconde :



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



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 :

$$\text{soit : } I_{gg} = 0$$

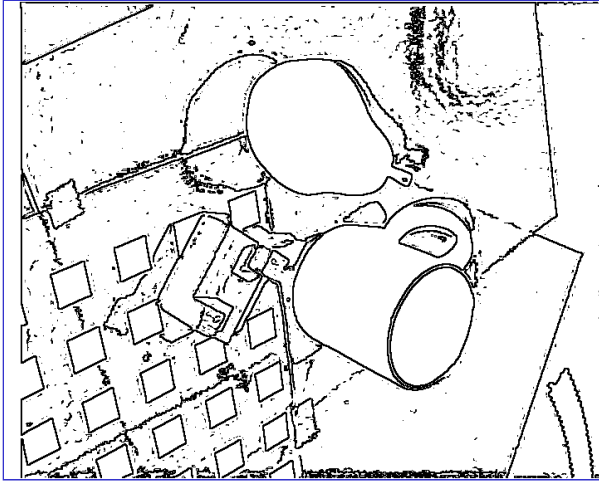
$$\text{or : } \begin{cases} I_g = I_x \cos \phi + I_y \sin \phi \\ I_t = I_x \sin \phi - I_y \cos \phi \end{cases} \quad \text{et : } \begin{cases} I_{gg} = I_{xx} \cos^2 \phi + 2I_{xy} \cos \phi \sin \phi + I_{yy} \sin^2 \phi \\ I_{tt} = I_{xx} \sin^2 \phi - 2I_{xy} \cos \phi \sin \phi + I_{yy} \cos^2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

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

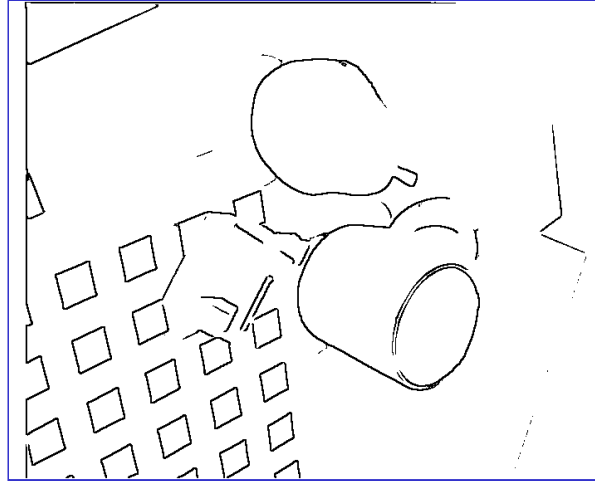


Passage par zéro du laplacien avec seuillage du module du gradient ($s = 1$) :

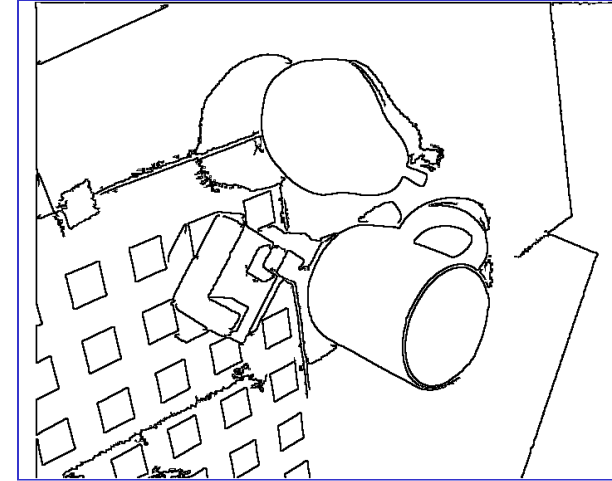
$$C_s(I) = \delta_B(\Delta_+ I) \cap \delta_B(\Delta_- I) \cap \sigma_s(\|\nabla I\|)$$

avec :

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



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]

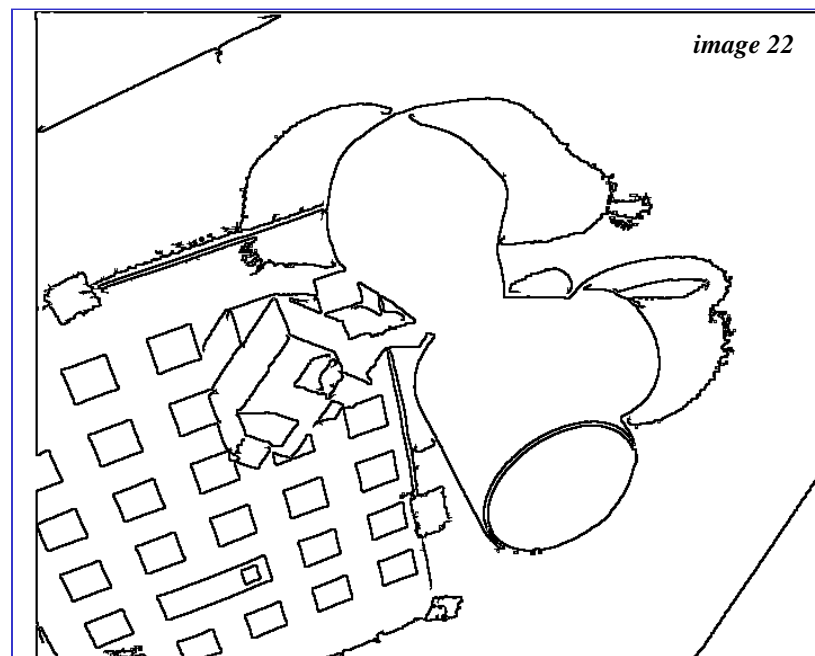
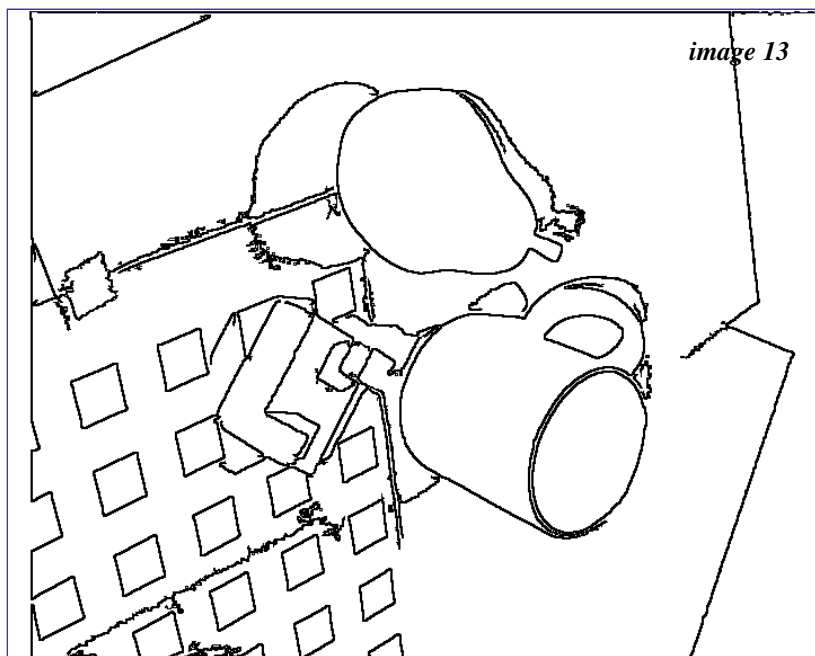
Détection de contours : applications

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 :
- Déplacement de la caméra
 - Orientation des facettes
 - Profondeur

Applications :

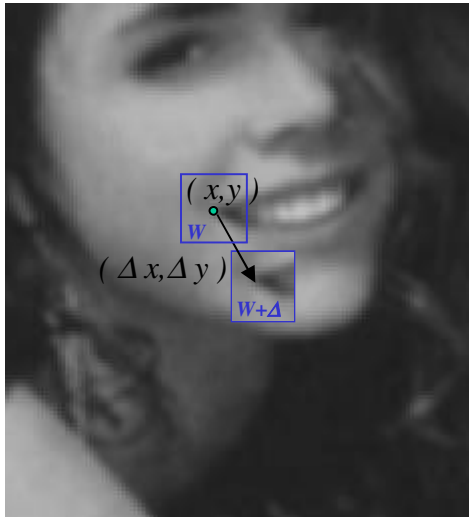
- Navigation / guidage de robots.
- Indexation d'images.



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 :

$$I(x_k + \Delta x, y_k + \Delta y) \approx I(x_k, y_k) + \left(\frac{\partial I}{\partial x}(x_k, y_k) \quad \frac{\partial I}{\partial y}(x_k, y_k) \right) \cdot \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix}$$

Et donc :

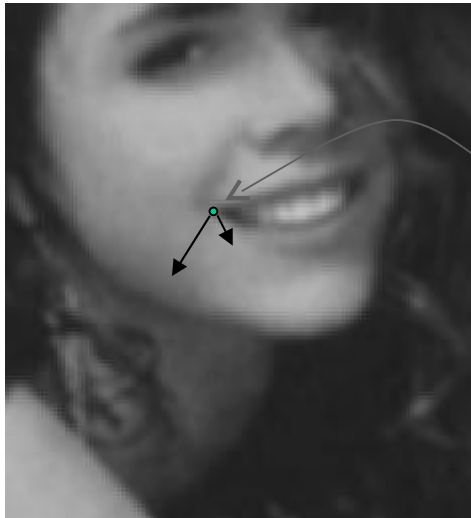
$$\chi(x, y) = \sum_{(x_k, y_k) \in W} \left(\left(\frac{\partial I}{\partial x}(x_k, y_k) \quad \frac{\partial I}{\partial y}(x_k, y_k) \right) \cdot \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \right)^2 = \begin{pmatrix} \Delta x & \Delta y \end{pmatrix} \underbrace{\begin{pmatrix} \sum_{(x_k,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

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 Ξ 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 Ξ 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.

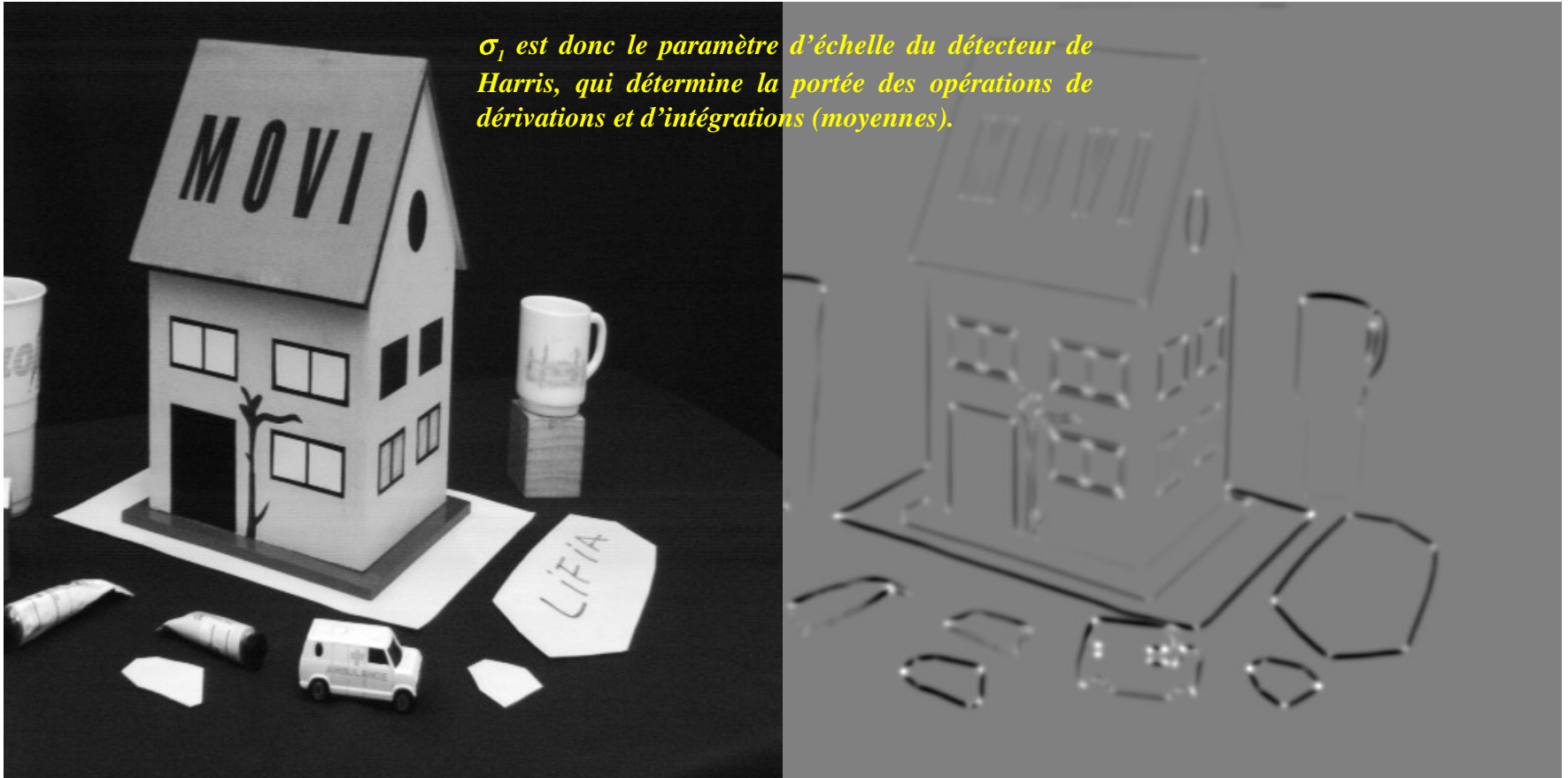
Les points d'intérêt correspondent aux maxima locaux de la fonction Θ qui sont au delà d'un certain seuil (typiquement 1% de la valeur max de Θ).

[Harris 88]

Mise en oeuvre du détecteur de Harris

1. On calcule les dérivées premières à partir des dérivées de gaussienne (écart-type σ_1)
2. On calcule les termes de la matrice d'autocorrélation Ξ en calculant une moyenne locale des dérivées sous la forme d'une gaussienne (écart-type σ_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 Θ supérieurs à un certain seuil (typiquement 1% de Θ_{\max}).

σ_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 :

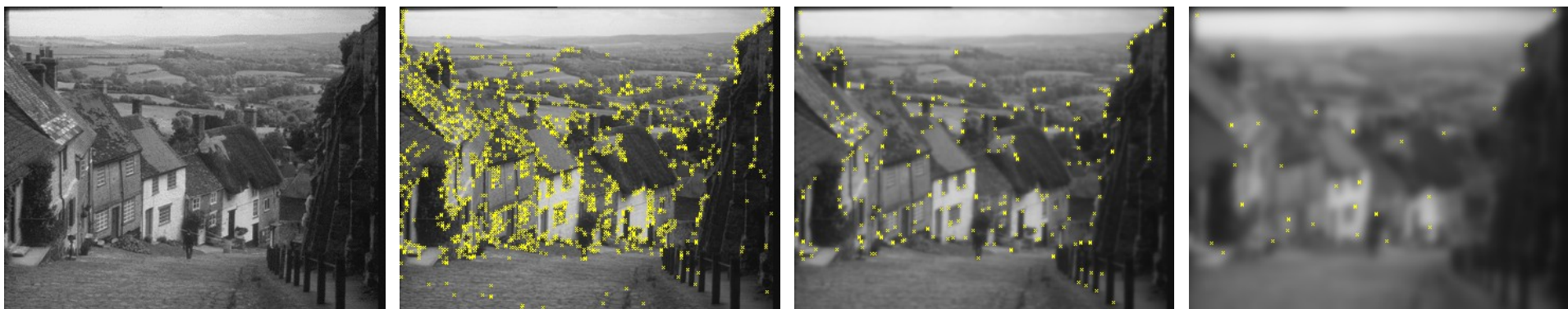


Applications : reconstruction, poursuite, indexation,...

Illustration du principe de causalité



Passage par zéro du laplacien (contours)



Détection de points d'intérêt (ou points anguleux)

Scale-space gaussien et diffusion

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

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

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

Diagram illustrating the components of the diffusion equation:

- $\frac{\partial I}{\partial t}$: gradient temporel
- div : divergence
- c : conductance
- ∇I : gradient spatial
- ΔI : 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).



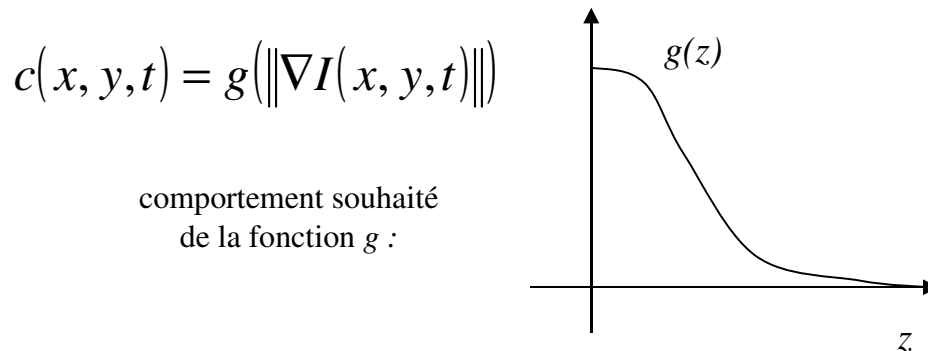
Espaces d'échelle
non linéaires

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 :



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

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

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 :

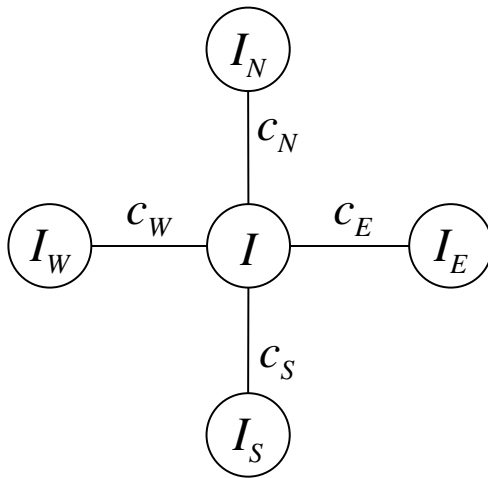
Le laplacien est discrétisé aux 4 plus proches voisins.

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

L'équation d'évolution discrète est la suivante :

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

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



avec :

$$\begin{cases} \Delta_N I = I_N - I \\ \Delta_E I = I_E - I \\ \Delta_S I = I_S - I \\ \Delta_W I = I_W - I \end{cases} \quad \text{et :} \quad \begin{cases} c_N = g(|\Delta_N I|) \\ c_E = g(|\Delta_E I|) \\ c_S = g(|\Delta_S I|) \\ c_W = g(|\Delta_W I|) \end{cases}$$

$g(x) = e^{-\left(\frac{x}{K}\right)^2}$
 ou
 $g(x) = \frac{1}{1 + \left(\frac{x}{K}\right)^2}$

Rq₁ : 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.

Rq₂ : 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.

Les filtres bilatéraux

Le schéma numérique précédent induit un calcul itératif, qui s'avère très coûteux lorsque l'étendue spatiale du filtre (l'échelle) devient importante. Un comportement équivalent de filtrage préservant les discontinuités peut être obtenu en utilisant les *filtres bilatéraux* :

Filtrage dans le domaine spatial :

$$g(x) = K_s^{-1} \iint f(y)h(y-x)dy \quad K_s = \iint h(y)dy$$

Filtrage dans le domaine tonal :

$$g(x) = K_t^{-1}(x) \iint f(y)s(f(y)-f(x))dy \quad K_t(x) = \iint s(f(y)-f(x))dy$$

Filtrage bilatéral :

$$g(x) = K^{-1}(x) \iint f(y)h(y-x)s(f(y)-f(x))dy \quad K(x) = \iint h(y-x)s(f(y)-f(x))dy$$

Exemple d'implantation gaussienne numérique :

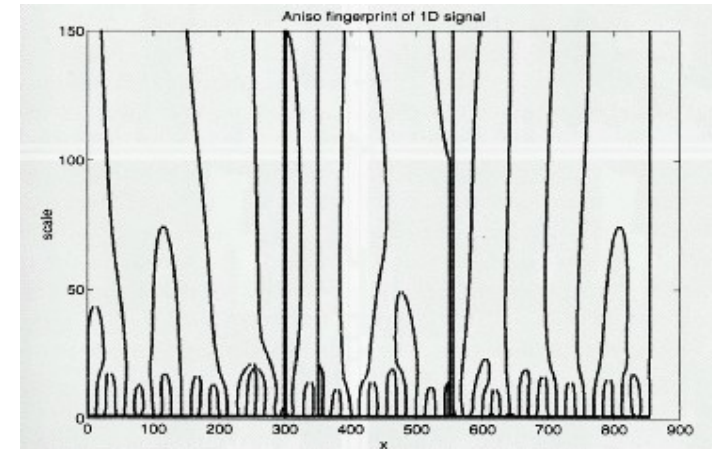
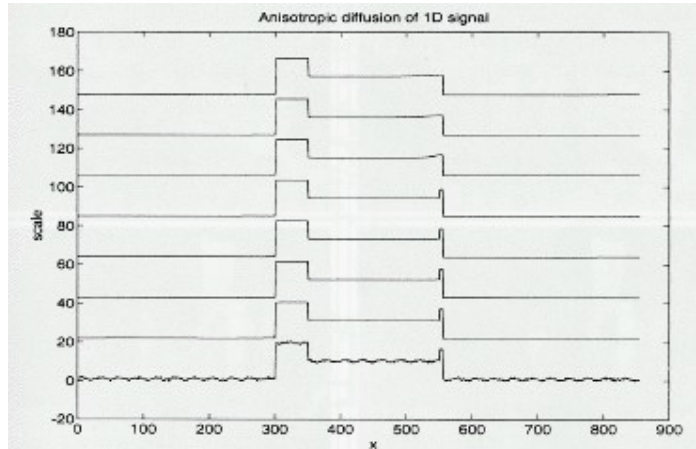
$$J[x] = \frac{\sum_{k=-N}^{+N} W[k, x]I[x-k]}{\sum_{k=-N}^{+N} W[k, x]}$$

avec : $W[k, x] = W_s[k] \cdot W_t[k, x]$

$$W_s[k] = e^{-\frac{k^2}{\sigma_s^2}} \quad W_t[k, x] = e^{-\frac{(I[x]-I[x-k])^2}{\sigma_t^2}}$$

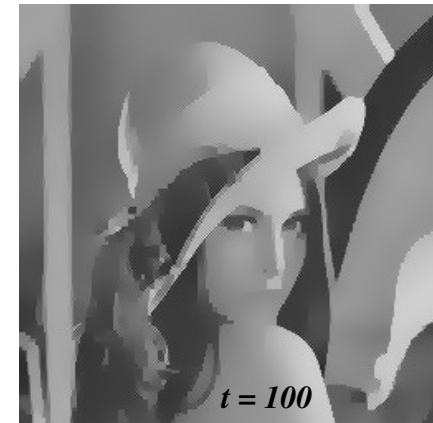
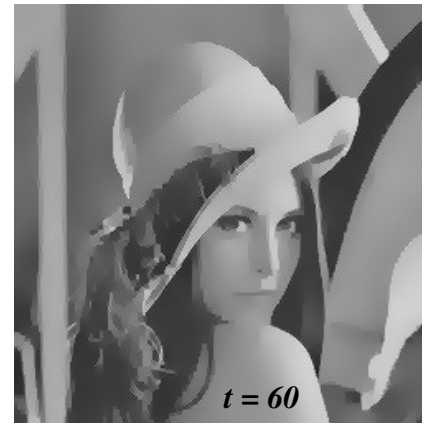
[Tomasi et Manduchi 98]

La pyramide de diffusion anisotrope



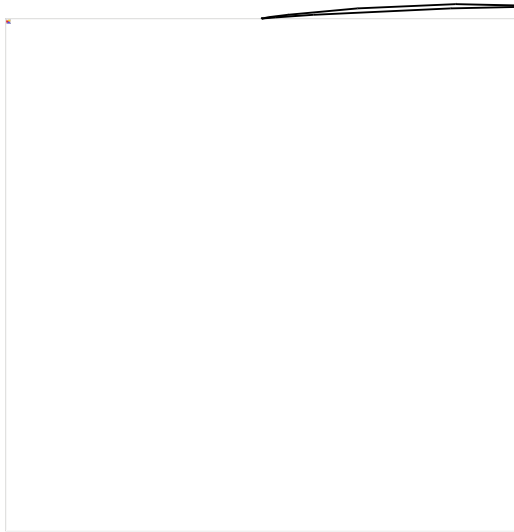
diffusion anisotrope d'un signal 1d

position des extrema dans l'espace d'échelles

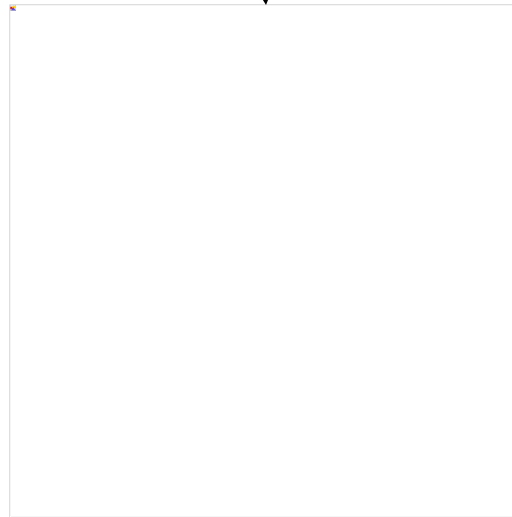


Diffusion anisotrope, schéma de décroissance hyperbolique (Image sur 8 bits, $K=15$).

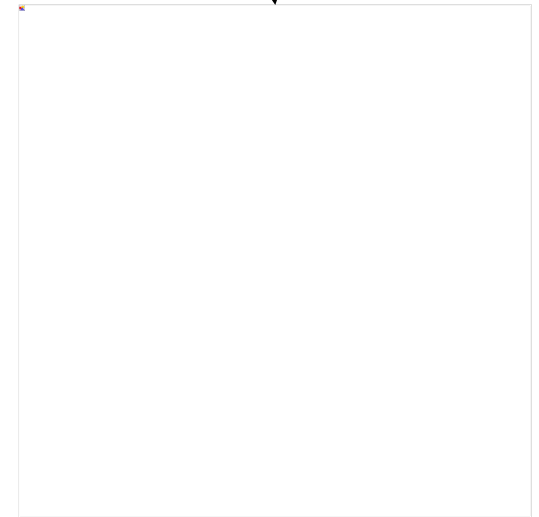
Diffusion anisotrope et extrema



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



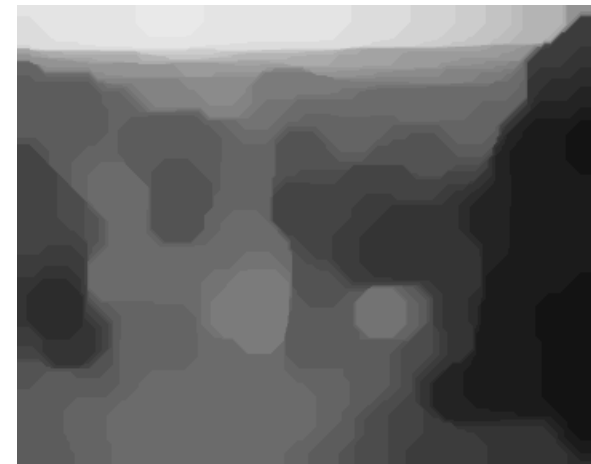
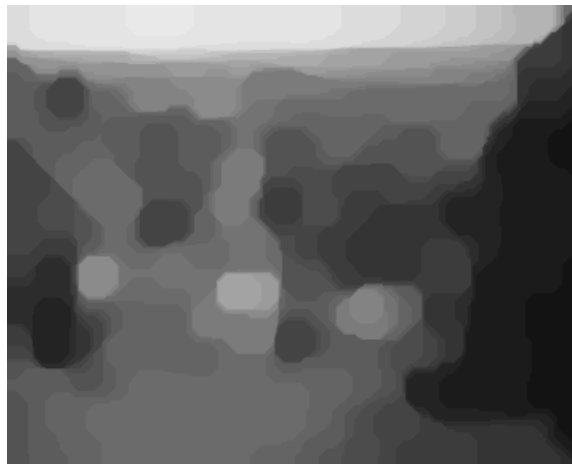
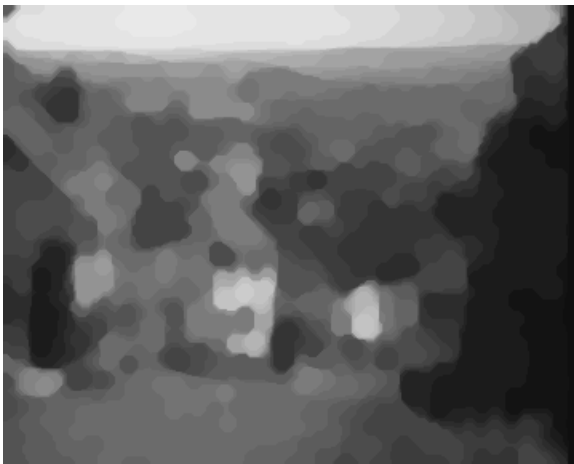
*Passage par zéro du loG ($\sigma = 1,5$)
avec seuil par hystérésis du module
du gradient ($S_{haut} = 5; S_{bas} = 1$).*



*Passage par zéro du loG ($\sigma = 1,5$)
avec seuil simple du module du
gradient ($S = 5$).*

Premier scale-space morphologique

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



Second scale-space morphologique

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



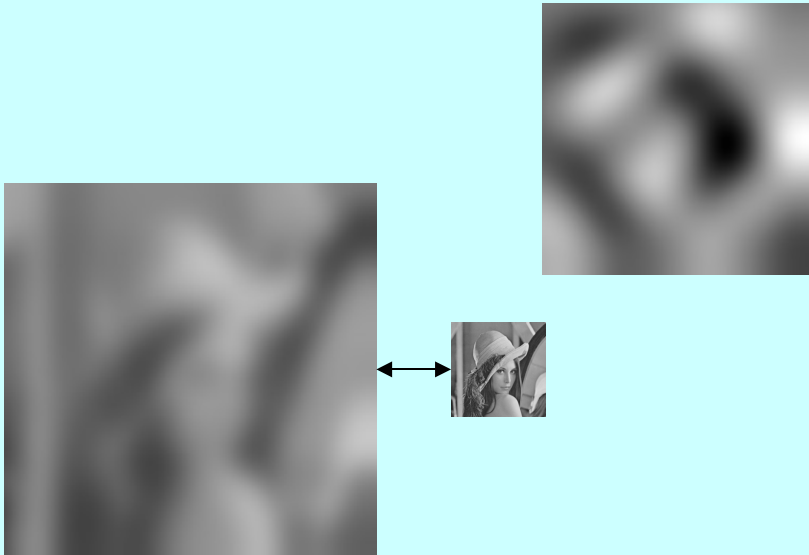
Bibliographie du chapitre

- **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*

Conclusion Scale-Space

A retenir pour ce cours :

- *Concept de Causalité*
- *Portée des opérateurs de TI*
- *Correspondance résolution / filtre*
- *Estimation des dérivées*



Applications : Contours / Points d'intérêt



Liens avec les EDP

