

# Formation et Analyse d'Images

James L. Crowley

ENSIMAG 3

Premier Sémestre 2006/2007

Séance 6

27 octobre 2006

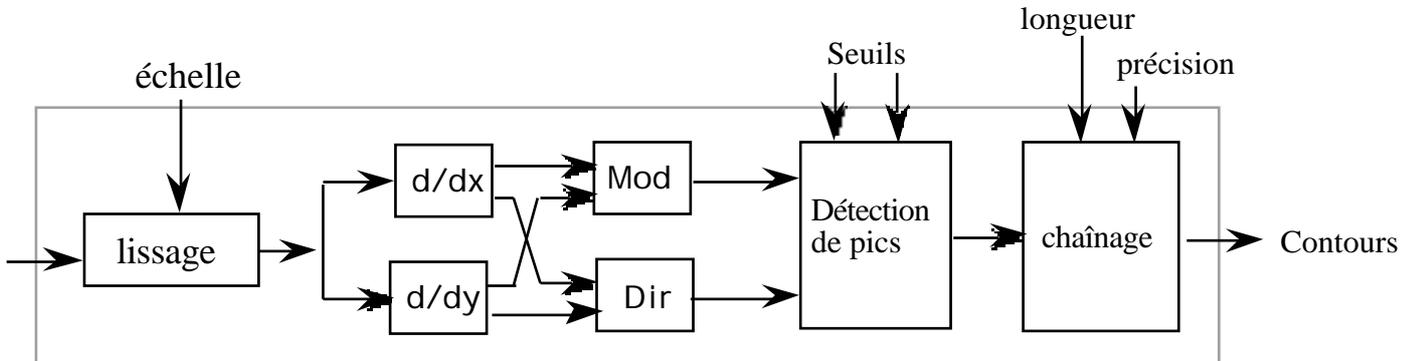
## Détection et Description de Contraste

### Plan de la Séance :

Détection de contrastes par dérivées.....	2
Les Champs Réceptifs Gaussiens .....	3
Filtres et dérivées.....	3
La fonction Gaussien.....	4
Les dérivées de la Gaussien.....	5
Les operateurs Differentielles.....	5
Les Filtres Numérique Gaussien.....	6
Les Dérivées de l'Image.....	8
L'Orientabilité des Gaussiens.....	8
L'espace d'échelles .....	9
Echelle Intrinsèque (ou caractéristique).....	10

## Détection de contrastes par dérivées

Schéma du processus de description de contraste



$$P(i, j) = \frac{\frac{p(i,j)}{i}}{\frac{p(i,j)}{j}} = \frac{m_0 * p(i,j)}{m_{90} * p(i,j)} = \frac{E_0(i,j)}{E_{90}(i,j)}$$

$$\text{Module} = \| P(i, j) \| = E(i, j) = \sqrt{E_1(i, j)^2 + E_2(i, j)^2}$$

$$\text{Direction} = (i, j) = \text{Tan}^{-1}\left(\frac{E_2(i, j)}{E_1(i, j)}\right)$$

### Étapes :

- 1) Lissage : suppression du bruit, conditionnement du signal
- 2) Calcul de gradient magnitude et orientation
- 3) Détection des maxima locaux avec seuillage par hystérésis.
- 4) Chaînage des points
- 5) Approximation polygonale : par aire, orientation, etc...

## Les Champs Réceptifs Gaussiens

### Filtres et dérivées

La produit scalaire est une projection d'un signal sur une fonction.

$$\langle f(n), s(m-n) \rangle = \sum_{n=-\infty}^{\infty} s(m-n) f(n)$$

Une evaluation de la produit scalaire à chaque pixel est un "convolution".

$$r(m) = f * s(m) = \sum_{n=-\infty}^{\infty} s(m-n) f(n) = \sum_{n=-\infty}^{\infty} s(n) f(m-n)$$

Pour calculer la dérivé d'un signal numérique, on calcule la fonction de la dérivé d'un noyau,  $g(t)$ . ensuite, on échantillon

$$g_x(x) = \frac{dg(x)}{dx}$$

$$\frac{ds(n)}{dt} = g_x(n-x) * s(n) \text{ avec } x=1$$

L'idéal serait une fonction invariante au tranformation projective. Mais ceci n'est pas possible. Mais on peut trouver une noyau invariante aux transformations affines - La fonction Gaussien

**La fonction Gaussien**

La fonction Gaussien est  $G(x, y) = e^{-\frac{1}{2} \frac{x^2+y^2}{z_c}}$

La fonction Gaussienne est invariante à la transformation affine:

$$T_a\{G(x, y)\} = G(T_a\{x\}, T_a\{y\})$$

Rappel en séance 2 on a vu que  $x_r = x_c \frac{F}{z_c}$ .

Donc la "taille" d'un objet est en proportion de  $s = \frac{F}{z_c}$ .

La taille (ou échelle) est une paramètre de la transformation affine.

La fonction Gaussienne est invariante à la transformation d'échelle :

$$T_s\{G(x, y)\} = G(T_s\{x\}, T_s\{y\}) = G(sx, sy)$$

Si on divise  $z_c$  (distance entre la caméra et l'objet) par deux, on double la taille.

$$G(x, y) = G(2x, 2y)$$

En 2-D, la Gaussienne symétrique circulaire est  $G(x, y) = e^{-\frac{x^2+y^2}{2A^2}}$  avec  $A = \frac{z_c}{2F}$

La Gaussien est la fonction unique qui est symétrique circulaire et séparable.

On note que  $G(x, y) = e^{-\frac{x^2+y^2}{2A^2}} = e^{-\frac{x^2}{2A^2}} * e^{-\frac{y^2}{2A^2}}$

Ceci nous offre beaucoup d'intérêt pour la vitesse de calcul.

## Les dérivées de la Gaussien

$$G(x, y) = e^{-\frac{1}{2} \frac{x^2 + y^2}{\sigma^2}}$$

$$G_x(x, y) = -\frac{x}{\sigma^2} G(x, y)$$

$$G_{xx}(x, y) = \left( \frac{x^2}{\sigma^4} - \frac{2}{\sigma^2} \right) G(x, y)$$

$$G_{xxx}(x, y) = -\frac{3x^3 - x}{\sigma^6} G(x, y)$$

On note que le "gain" d'un noyau est son intégral.

$$A = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{1}{2} \frac{x^2 + y^2}{\sigma^2}} dx dy = \sqrt{2\pi} \sigma^2$$

La propriété de l'invariance à l'échelle demande que  $A = 1$ .

## Les opérateurs Différentielles

Pour un signal en deux dimensions, les opérations différentielles sont le gradient et la Laplacien <sup>2</sup> :

Le gradient est un vecteur :

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{pmatrix}$$

La Laplacien est une scalaire :  $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$

Pour une fonction  $s(x, y)$  :

La Gradient :

$$\nabla s(x, y) = \begin{pmatrix} \frac{\partial s(x, y)}{\partial x} \\ \frac{\partial s(x, y)}{\partial y} \end{pmatrix}$$

La Laplacien

$$\nabla^2 s(x, y) = \frac{\partial^2 s(x, y)}{\partial x^2} + \frac{\partial^2 s(x, y)}{\partial y^2}$$

si  $s(x,y)$  est filtré par  $G(x,y; \sigma)$  :  $s * G(x,y; \sigma) = \int_{s=-\infty}^{\infty} \int_{t=-\infty}^{\infty} s(x-s, y-t) G(x-s, y-t) dx dt$

La Gradient :  $(s * G(x,y; \sigma))_x = s * \frac{\partial G(x,y; \sigma)}{\partial x}$   
 $(s * G(x,y; \sigma))_y = s * \frac{\partial G(x,y; \sigma)}{\partial y}$

La Laplacien  $\Delta(s * G(x,y; \sigma)) = s * \Delta G(x,y; \sigma) = s * \left( \frac{\partial^2 G(x,y; \sigma)}{\partial x^2} + \frac{\partial^2 G(x,y; \sigma)}{\partial y^2} \right)$

### Les Filtres Numérique Gaussien

Parce que l'image est un signal échantillonné, il faut échantillonner  $G(x, y, \sigma)$

On obtient les filtres numériques par un simple échantillonnage de la fonction Gaussienne sur un intervalle  $[-R, R]$ .

On remplace  $x$  par  $nT_e$ , ou  $T_e$  est un

$$G(n) = G(nT_e, \sigma) = e^{-\frac{1}{2} \frac{n^2 T_e^2}{\sigma^2}}$$

$T_e$  est la pas d'échantillonnage.

Par convention l'on considère  $T_e = 1$ .

Donc, la forme numérique est  $G(n, \sigma) = e^{-\frac{1}{2} \frac{n^2}{\sigma^2}}$

$$A = \int_{n=-R}^R e^{-\frac{1}{2} \frac{n^2}{\sigma^2}} dx \sqrt{2\pi} \sigma$$

Il y a deux facteurs à maîtriser :

- a) La taille de la "support"  $N = 2R+1$ .
- b) la ratio  $\sigma/T_e$

Pour a:

Pour  $N = 7$ , les "ondes" de  $W_R(f)$  dominant le spectre.

Pour  $N = 9$ , les ondes, on peut d'effet.

Pour b:

Il vaut mieux que  $\sigma/T_e = 1$

Les dérivées de la Gaussienne numérique sont :

$$G_x(n, ) = -\frac{n}{2} G(n, )$$

$$G_{xx}(n, ) = \frac{n^2 - 2}{4} G(n, )$$

$$G_{xxx}(n, ) = -\frac{n^3 - n}{6} G(n, )$$

Pour la Gaussien en 2D.  $G_x(i, j, )$

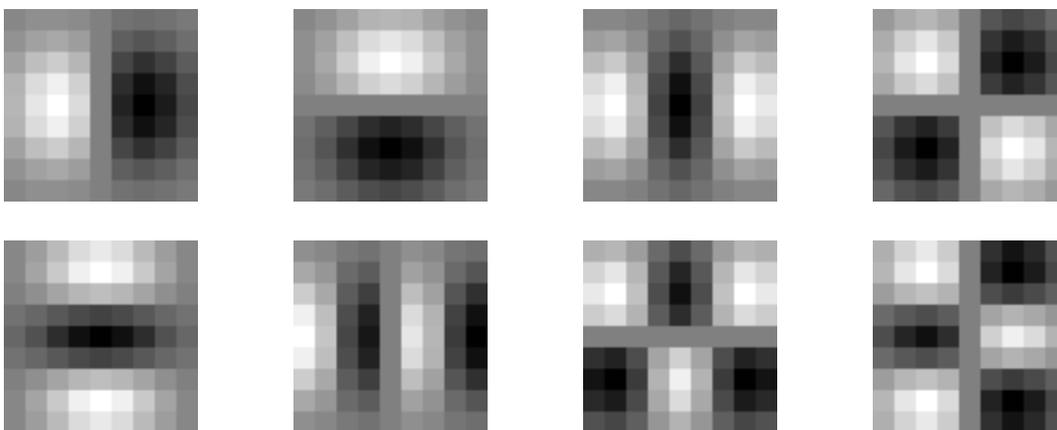
$$G_x(i, j, ) = -\frac{i}{2} G(i, j, ) = -\frac{i}{2} G(i, ) * G(j, )$$

$$G_{xy}(i, j, ) = \frac{i}{2} \frac{j}{2} G(i, j, ) = -\frac{i}{2} G(i, ) * -\frac{j}{2} G(j, )$$

Un vecteur de champs réceptifs forme une base de Taylor

$$G_a = (G_x, G_y, G_{xx}, G_{xy}, G_{yy}, G_{xxx}, G_{xxy}, G_{xyy}, G_{yyy})$$

Ceci donne la famille de champs réceptifs Gaussien



Les champs réceptifs Gaussien  $G_x, G_y, G_{xx}, G_{xy}, G_{yy}, G_{xxx}, G_{xxy}, G_{xyy}, G_{yyy}$ .

Note qu'il y a une paramètre  $\sigma$ . Ceci est la paramètre d'échelle. Ce détermine la limite de la résolution d'une description.

## Les Dérivées de l'Image

Pour l'image  $p(i, j)$ ,  $p(i, j)$  est calculé par  $G(i, j, \theta) * p(i, j)$ .

$$\text{ou } G(i, j, \theta) = \begin{pmatrix} G_x(i, j, \theta) \\ G_y(i, j, \theta) \end{pmatrix}$$

$$\text{Gradient: } \nabla p(x, y) = \begin{pmatrix} G_x(x, y, \theta) * p(x, y) \\ G_y(x, y, \theta) * p(x, y) \end{pmatrix}$$

(Le gradient est un vecteur).

$$\text{Laplacien: } \nabla^2 p(i, j) = G_{xx}(i, j, \theta) * p(i, j) + G_{yy}(i, j, \theta) * p(i, j)$$

(Le Laplacien est une scalaire.)

## L'Orientabilité des Gaussiens

Pour chaque pixel, on peut calculer une orientation "intrinsic"

$$\theta(x, y) = \text{Atan2}\{\langle A(x, y), G_y \rangle, \langle A(x, y), G_x \rangle\}$$

Les réponses de filtres orientées peuvent être calculé par une somme de réponse des filtres de base, pondérée par les sinus et cosinus.

$$G_1^\theta = \cos(\theta) G_x + \sin(\theta) G_y$$

$$G_2^\theta = \cos(\theta)^2 G_{xx} + 2 \cos(\theta) \sin(\theta) G_{xy} + \sin(\theta)^2 G_{yy}$$

$$G_3^\theta = \cos(\theta)^3 G_{xxx} + 3 \cos(\theta)^2 \sin(\theta) G_{xxy} + 3 \cos(\theta) \sin(\theta)^2 G_{xyy} + \sin(\theta)^3 G_{yyy}$$

## L'espace d'échelles

La fonction Gaussian est invariante à la transformation d'échelle :

$$T_s\{G(x, y)\} = G(T_s\{x\}, T_s\{y\})$$

On a vu que la "taille" d'un objet est en proportion de  $\frac{F}{z_c}$ .

Si on multiplie par "s" la taille.

$$G(sx, sy) = G(x, y)$$

Mais, on concrète, quelle échelle faut-il utiliser ? tous !

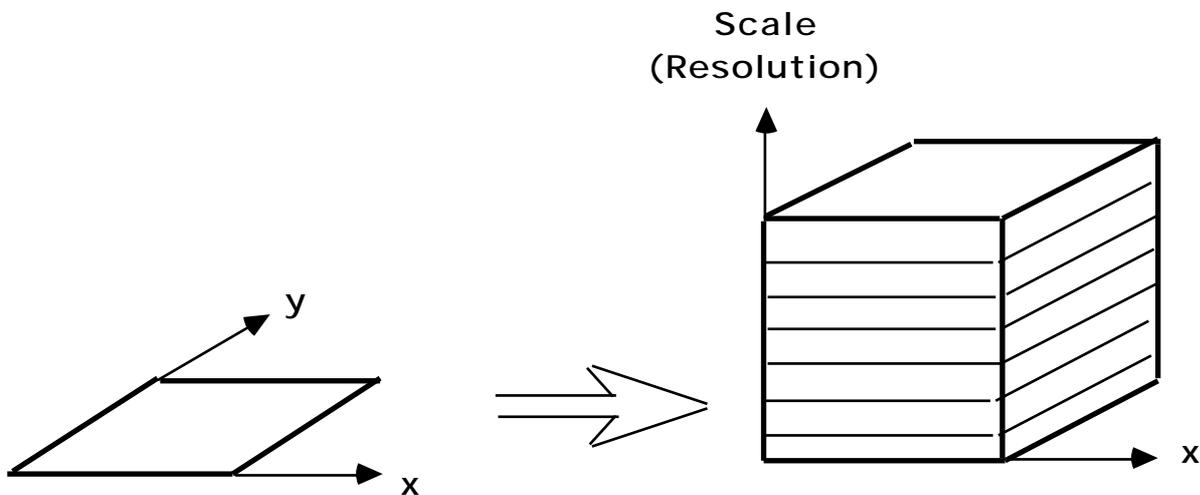
Un espace d'échelle,  $P(x, y, s)$  est défini par un noyau,  $g(x, y; s)$  avec le paramètre  $s$  libre

$$P(x, y, s) = g(x, y, s) * p(x, y)$$

pour  $X_{\min} \leq x \leq X_{\max}$ ,  $Y_{\min} \leq y \leq Y_{\max}$ ,  $s_{\min} \leq s \leq s_{\max}$ ,

ou  $s$  est le paramètre d'échelle.

Pour une description invariant à l'échelle, il faut un axe logarithmique pour  $s$ .



La description de l'image se trouvent à toutes les échelles.

L'espace d'échelle facilite la recherche de correspondance grâce à la décomposition des formes à divers niveaux de résolution :

- |                    |   |
|--------------------|---|
| Résolution basse : | peu de formes,<br>vue globale,<br>détails grossiers |
| Résolution haute : | nombreux formes,<br>vue locale,<br>détails fins.    |

Propriétés de l'espace d'échelle :

- Invariance (Equivariance) aux changements de taille
- Le bruit de numérisation se trouve dans les hautes fréquences (donc petit  $\sigma$ ).
- Les points de contraste dans les moyennes fréquences sont souvent les plus stables,

L'espace d'échelle est un idéal mathématique. Pour calculer il faut travailler sur les valeurs numériques. Ceci implique un échantillonnage en  $x$ ,  $y$ , et  $\sigma$ .

Le Gaussian est une solution de l'Equation de Diffusion.

Equation de Diffusion:  $\frac{\partial^2 G(x,y; \sigma)}{\partial x^2} = -\frac{G(x,y; \sigma)}{\sigma^2}$

En conséquence,  $\frac{\partial^2 G(x,y; \sigma)}{\partial x^2} = G(x,y; \sigma_1) - G(x,y; \sigma_2)$

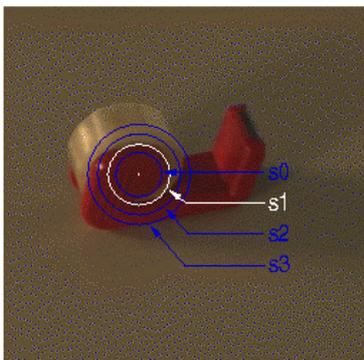
pour  $\sigma_1 = \sqrt{2} \sigma_2$

### Echelle Intrinsèque (ou caractéristique)

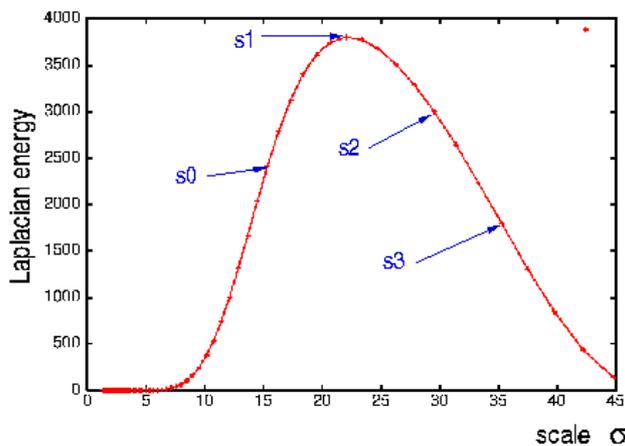
Considère la Laplacienne pour à la pixel  $x, y$ , en fonction de

$$\Delta^2 p(i,j) = \langle G_{xx}(i, j, \sigma), p(i, j) \rangle + \langle G_{yy}(i, j, \sigma), p(i, j) \rangle$$

A chaque point de l'image il y a quelques valeurs de  $\sigma$  pour laquelle la  $\Delta^2 p(i,j)$  sont maximales.



zero crossing of Laplacian at  $\sigma_i$



L'echelle intrinsèque à  $(i, j)$  est  $\sigma_i(i, j) = \text{Arg-Max} \{ \Delta^2 p(i,j) \}$

Les dérivés donnent les caractéristiques locale. Par normalisation à l'échelle et orientation ils deviennent invariants

$$X(i, j) = p(i, j) * \begin{matrix} G_x \\ G_y \\ G_{xx} \\ G_{xy} \\ G_{yy} \end{matrix}$$

Pour chaque classe k, la probabilité d'observer X à le pixel i, j, est fournie par la règle de Bayes.

$$p(k | X) = \frac{p(X | k) p(k)}{p(X)}$$

On peut transformer  $X_d$  en entier entre [0, N-1], par

$$X_1 = \langle p(i, j), G_x \rangle = \text{Round} \left( N \cdot \frac{\langle p(i, j), G_x \rangle}{V_{\max}} \right).$$

Soit N bins par dimension, avec D dimensions.

La valeur de N doit être choisi pour assurer que  $Q = 2^{ND} < 10^6$  observations.

Dans ce cas, pour M observations  $p(X) = \frac{1}{M} h(X)$

La probabilité à posteriori peut être calculé par la règle de Bayes.

soit  $k$ , la proposition que l'observation X la classe k

$$p(k | X) = \frac{p(X | k) p(k)}{p(X)}$$

Dans le cas des valeurs de X discrètes tel que  $x \in [X_{\min}, X_{\max}]$ , on a

probabilité de la classe k:  $p(k) = \frac{M_k}{M}$

probabilité conditionnelle de X :  $p(X | k) = \frac{1}{M_k} h_k(x)$

Probabilité à priori de  $X$  :  $p(X) = \frac{1}{M} h(x)$

ce qui donne :

$$p(k | X) = \frac{p(X | k) p(k)}{p(x)} = \frac{\frac{M_k}{M} \frac{1}{M_k} h_k(x)}{\frac{1}{M} h(x)} = \frac{h_k(x)}{h(x)}$$