Formation et Analyse d'Images

James L. Crowley

ENSIMAG 3 — Premier Sémestre 2007/2008

Séance 6 — 9 novembre 2007

Détection et Description de Contraste

Plan de la Séance :

Les Champs Réceptifs Gaussiens

Filtres et dérivées

La Transformé de Fourier d'une dérivée a une fonction est

$$\mathcal{F}\left\{\frac{\partial s(t)}{\partial t}\right\} = -2\pi j\, \omega\, \mathcal{F}\{f(t)\}$$

et donc $\frac{\partial s(t)}{\partial t} = \mathcal{F}^{-1}\{-2\pi j\, \omega\, f(t)\} = \mathcal{F}^{-1}\{-2\pi j\, \omega\} * \mathcal{F}^{-1}\{f(t)\}$

Donc, une dérivée est un FILTRE avec une fonction de transfert $-2\pi j\, \omega$

$$\frac{\partial f(t)}{\partial t} = \frac{\partial}{\partial t} * s(t) = \mathcal{F}^{-1}\{-2\pi j\, \omega\} * s(t)$$

Les filtres linéaires sont associatifs, distributifs et commutative.

$\mathcal{F}^{-1}\{-2\pi j\, \omega\}$ a une durée infinie. Mais on peut faire une approximation de durée finie par

$$d(n) = [-1, 0, 1]$$

En alternative, on peut calculer la fonction de la dérivéed'un noyau, g(t). ensuite, on échantillon

$$f(t) = \frac{dg(t)}{dt}$$

$$\frac{ds(n)}{dt} = f(nT_e) * s(n)$$

On note que le convolution est une produit scalaire evalué à chaque echantillion

$$r(k) = s * f(i) = <s(i), f(k-i)> = \sum_{i=-\infty}^{\infty} s(k-i)\, f(i) = \sum_{i=-\infty}^{\infty} s(i)\, f(k-i)$$

Autrement dit, le produit scalaire est une projection d'un signal sur une fonction.

$$<s, f(k)> = \sum_{i=-\infty}^{\infty} s(i)\, f(i+k)$$

L'idéel 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 $G(x, \sigma) = e^{-\frac{1}{2}\frac{x^2}{\sigma^2}}$

La fonction Gaussienne est invariante à la transformation affine:

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

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

Donc la "taille" d'un objet et 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 :

$$T_s\{G(x, \sigma)\} = G(T_s\{x\}, T_s\{\sigma\}) = G(sx, s\sigma)$$

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

$$G(x, \sigma) = G(2x, 2\sigma)$$

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

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

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

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

Les dérivées de la Gaussien

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

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

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

$$G_{xxx}(x, \sigma) = -\frac{x^3 - x\,\sigma^2}{\sigma^6} \quad G(x, \sigma)$$

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

$$A = \int_{-\infty}^{\infty} e^{-\frac{1}{2}\frac{x^2}{\sigma^2}} \, dx = \sqrt{2\pi} \;\; \sigma .$$

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

Les operateurs Differentielles

Pour un signal en deux dimensions, les opérations différentielles sont le gradient $\vec{\nabla}$ et la Laplacien ∇^2 :

Le gradient est un vecteur : $\vec{\nabla} = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{pmatrix}$

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

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

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

La Gradient : $\vec{\ }(s * G(x,y;\sigma)) = s * \vec{\ }G(x,y;\sigma) = \begin{pmatrix} s * \dfrac{\partial G(x,y;\sigma)}{\partial x} \\ s * \dfrac{\partial G(x,y;\sigma)}{\partial y} \end{pmatrix}$

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

Les Filtres Numérique Gaussien

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

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

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

T_e est la pas d'échantillonnage.
Par convention l'on considère $T_e = 1$.

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

$$A = \sum_{n=-\infty}^{\infty} e^{-\frac{1}{2}\frac{n^2}{\sigma^2}} \, dx \approx \sqrt{2\pi}\ \sigma .$$

Il y a deux facteurs à maîtriser :

a) La taille de la "support” N = 2R+1.
b) la ratio σ/T_e

Pour a:
Pour N ≤ 7, les "ondes" de $W_R(f)$ dominent le spectre.
Pour N ≥ 9, les ondes, on peut d'effet.

Pour b:
Il vaut mieux que $\sigma/T_e \geq 1$

Les dérivées de la Gaussienne numérique sont :

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

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

$$G_{xxx}(n, \sigma) = -\frac{n^3 - n\,\sigma^2}{\sigma^6} \; G(n, \sigma)$$

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

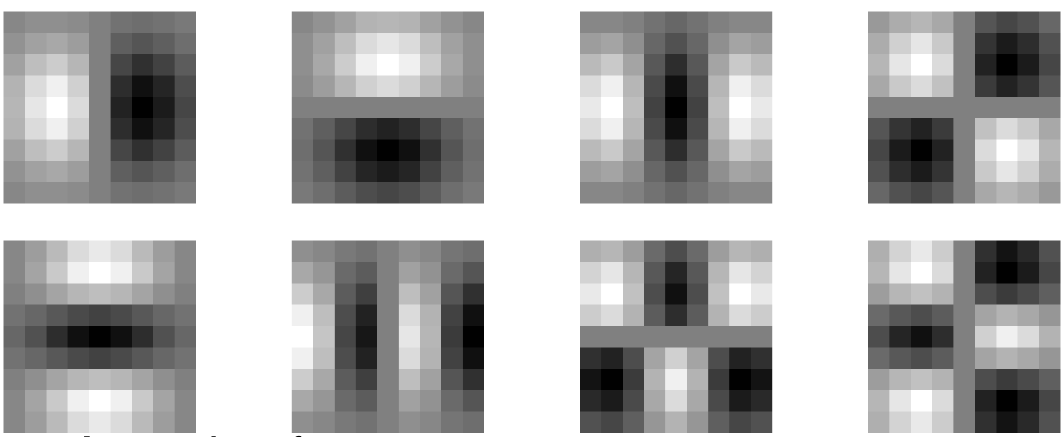
$$G_x(i, j, \sigma) = -\frac{i}{\sigma^2} \; G(i, j, \sigma) = -\frac{i}{\sigma^2} \; G(i, \sigma) * G(j, \sigma)$$

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

Un vecteur de champs réceptifs forme une base de Taylor

$$\vec{G}_a = (G_x, G_y, G_{xx}, G_{xy}, G_{yy}, G_{xxx}, G_{xxy}, G_{xyy}, G_{yyy})$$

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 σ. Ceci est la paramètre d'echelle. Ce détermine la limite de la résolution d'une déscription.

Les Dérivées de l'Image

Pour l'image p(m, n), $\vec{\nabla} p(m,n)$ est calculé par $\vec{\nabla} G(m, n, \sigma) * p(m,n)$.

ou $\vec{\nabla} G(m, n, \sigma) = \begin{pmatrix} G_x(m, n, \sigma) \\ G_y(m, n, \sigma) \end{pmatrix}$

Gradient: $p(x,y) = \begin{bmatrix} G_x(x, y, \sigma) * p(x, y) \\ G_y(x, y, \sigma) * p(x, y) \end{bmatrix} \quad \begin{bmatrix} G_x(x, y, \sigma) * p(x, y) \\ G_y(x, y, \sigma) * p(x, y) \end{bmatrix}$

(Le gradient est un vecteur).

Laplacien: $\nabla^2 p(m,n) = G_{xx}(m, n, \sigma) * p(m, n) + G_{yy}(m, n, \sigma) * p(m\ n)$

(La Laplacien est une scalaire.)

L'Orientabilité des Gaussiens

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

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

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 :

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

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

Si on multiplie par "s" la taille.

$$G(sx, s\sigma) = G(x, \sigma)$$

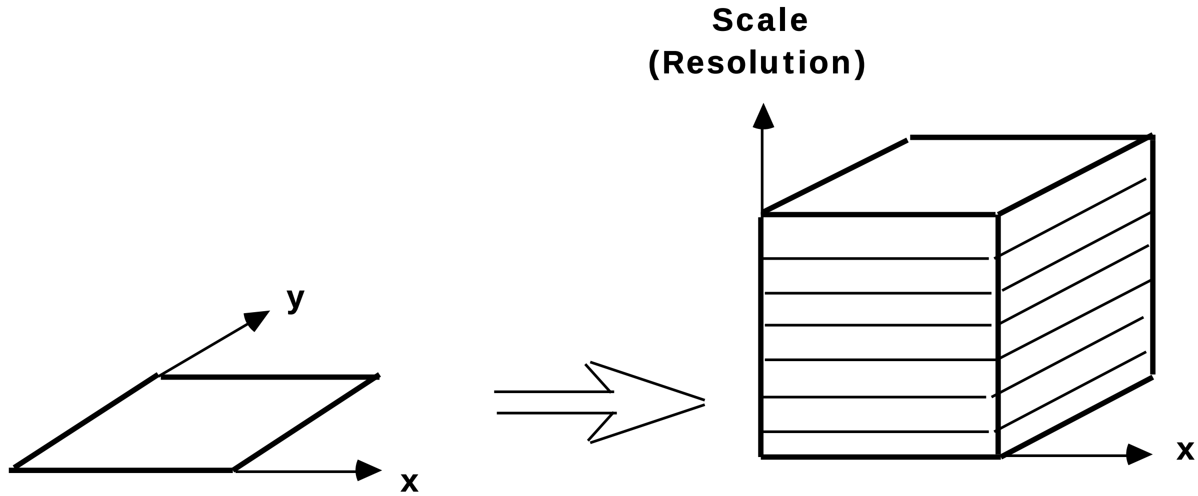
Mais, on concrète, quelle échelle faut-il utiliser ? tous !

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

$$P(x, y, \sigma) = g(x, y, \sigma) * p(x, y)$$

pour $X_{min} \leq x \leq X_{max}$, $Y_{min} \leq y \leq Y_{max}$, $\sigma_{min} \leq \sigma \leq \sigma_{max}$,

ou σ est le paramètre d'échelle.
Pour une description invariant à l'échelle , il faut un axe logarithmique pour σ.



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,
vue globale,
détails grossiers
Résolution haute : nombreux formes,
vue locale,
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 σ).
- Les points de contraste dans les moyennes fréquences sont souvent les plus stables,

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

Le Gaussian est une solution de la Equation de Diffusion.

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

En conséquente, $^2G(x,y;\sigma) \approx G(x,y;\sigma_1) - G(x,y;\sigma_2)$

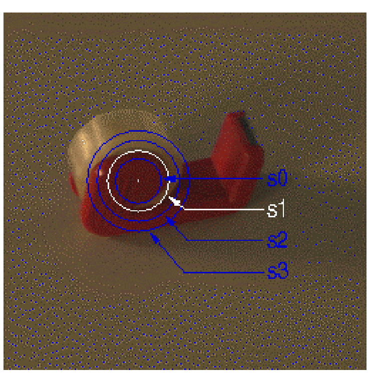
pour $\sigma_1 \geq \sqrt{2}\,\sigma_2$

Echelle Intrinsèque (ou characteristique)

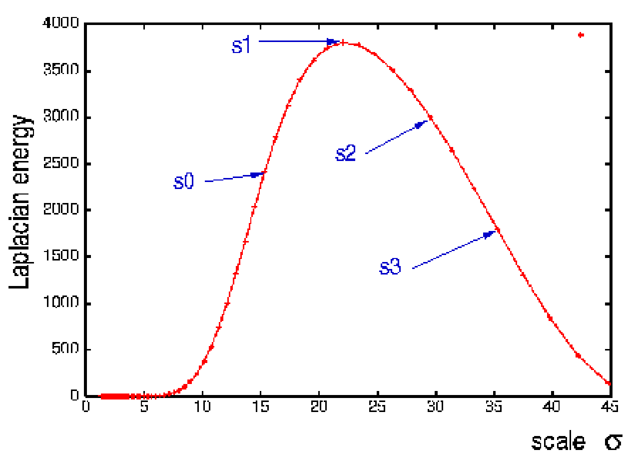
Considère la Laplacienne pour à la pixel x, y, en fonction de σ

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

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



zero crossing of Laplacian at si



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

Les dérivés donnes les caractéristiques locale. Par normalisation à l'echelle et orientation ils deviens invariantes

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

Pour chaque classe k, la probability d'observé $\vec{X}$ à le pixel i, j, est fournie par la règle de Bayes.

$$p(\omega_k \mid \vec{X}) = \frac{p(\vec{X} \mid \omega_k)\, p(\omega_k)}{p(\vec{X})}$$

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 $Q = 2^{ND} < 10$ M observations.

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

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

soit ω_k, la proposition que l'observation $\vec{X}$ la classe k

$$p(\omega_k \mid \vec{X}) = \frac{p(\vec{X} \mid \omega_k)\, p(\omega_k)}{p(\vec{X})}$$

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

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

probabilité conditionnelle de $\vec{X}$: $p(\vec{X} \mid \omega_k) = \frac{1}{M_k}\, h_k(\vec{x})$

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

ce qui donne :

$$p(\omega_k \mid \vec{X}) = \frac{p(\vec{X} \mid \omega_k)\, p(\omega_k)}{p(\vec{x})} \approx \frac{\frac{M_k}{M}\,\frac{1}{M_k}\; h_k(\vec{x})}{\frac{1}{M}\; h(\vec{x})} = \frac{h_k(\vec{x})}{h(\vec{x})}$$