Vision par Ordinateur

James L. Crowley

DEA IVR — Premier Bimestre 2005/2006

Séance 7 — 23 novembre 2005

Champs receptifs et l'espace d'echelles

Plan de la Séance :

Charactéristiques Locale

Objectif : Trouver les charactéristiques, $\vec{X}(i,j)$ pour les régions local d'une image

En séance 3 on a vue que la couleur est une charactéristique pour chaque pixel.

$$\vec{X}(i,j) = \begin{pmatrix} L \\ C_1 \\ C_2 \end{pmatrix}(i,j).$$

On a besoins des charactéristiques pour les voisinage de pixels.

Soit une image p(i,j).

$$\vec{V}(i,j) = \langle \vec{G}, p(i,j) \rangle \quad \text{est un vecteur de caractéristiques pour } i,j$$

Avec $\vec{V}(i,j)$ on peut faire les reconnaissance probabilistes. Pour l'objet O :

$$p(O(i,j) \mid \vec{V}(i,j)) = \frac{p(\vec{V}(i,j) \mid O(i,j))\, p(O(i,j))}{p(\vec{V}(i,j))} \approx \frac{h_o(\vec{V}(i,j))}{h_{tot}(\vec{V}(i,j))} = T_{ratio}(\vec{V}(i,j))$$

Ou bien on peut comparer une voisinage $\vec{V}$ à une modèle $\vec{M}$:

$$d^2 = \| \vec{V} - \vec{M} \|$$

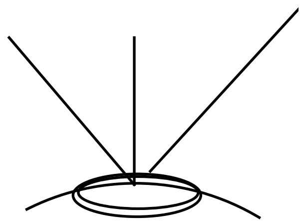
On peut également utiliser $\vec{V}(i,j)$ comme une cléf pour l'indexation.

Mais quel fonction $\vec{G}$ faut-il?

Détection et Description de Contraste

Rappelons que l'albédo d'un objet non-métalique peut être approximé par la composition d'une réflexion "spéculaire" et d'une réflexion "lambertienne".

$$R(i, e, g, \lambda) = \rho_S R_S(i, e, g, \lambda) + \rho_l(\lambda) R_l(i, e, g)$$



Le composant lambertien dépend de l'angle entre la normale et la source.

$$R_l(i, e, g) = \rho \operatorname{Cos}(i)$$

La composant "luminance" est déterminé par l'orientation de la surface.

Donc, les variations dans l'orientation d'une surface donne les variations dans la luminance

Soit $\alpha(i,j)$ l'angle d'incidence de la surface dans la scène qui est projetée à la pixel (i,j) de l'image p(i,j).Pour une déplacement s = $(\Delta i, \Delta j)$

$$\frac{\partial p(i,j)}{\partial s} = \frac{\partial \cos(\alpha(i,j))}{\partial s}$$

Donc les dérivées de p(i,j) décrivent la forme de la surface dans la scène.

Il convient de trouver une caractéristique dépendant de la forme locale.

Description par dérivées d'une fonction Gaussienne

La surface locale d'une image peut être approximé par suite de Taylor

$$s(t) \approx s(t_o) + \Delta t \frac{ds(t)}{dt} + \Delta t^2 \frac{d^2s(t)}{dt^2} + \Delta t^3 \frac{d^3s(t)}{dt^3} + \ldots$$

La suite de Taylor est une projection d'un signal sur un base composée de dérivées. Les coéficients de la projection forme une charactéristique locale, $\vec{v}(t)$

Pour une fonction, f(t)

$$\frac{df(t)}{dt} \equiv \lim_{\Delta t \to 0} \left\{ \frac{f(t) - f(t-\Delta t)}{\Delta t} \right\}$$

Pour un signal numérique, s(n), la limite n'existe pas.

$$\Delta n = 2 : \quad \frac{\partial s(n)}{\partial n} = \frac{s(n) - s(n-2)}{2}$$

$$\Delta n = 1 : \quad \frac{\partial s(n)}{\partial n} = \frac{s(n) - s(n-1)}{1}$$

$$\Delta n = 0 : \quad \frac{\partial s(n)}{\partial n} = \frac{0}{0} \quad !!!$$

Peut-on calculer une dérivée pour un signal discret ? oui. Comment faire ?

On calcule la produit local avec la dérivée d'une fonction "noyau", g(t).

< s(t), g(t)> est l'integral du produit sur une intervalle T.

$$< s(t), g(t)> = \int_{-T/2}^{T/2} s(\tau)\ g(\tau)\ d\tau$$

$$\frac{d(<s(t), g(t)>)}{dt} = <\frac{ds(t)}{dt}, g(t)> = <\frac{dg(t)}{dt}, s(t)>$$

Dans le cas ou s(t) est numérique s(n) = s(n T_e) avec pas d'échantillonnage T_e

Note que une produit calculé à tout les pixels s'appelle une convolution.

$$h(t) = s * g(t) = \int_{-T/2}^{T/2} s(\tau)\ g(t-\tau)\ d\tau$$

Filtres et dérivées

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

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

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

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

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

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

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

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

Les Champs Réceptifs Gaussiens

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.

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

Les dérivées de la Gaussien

Pour une suite de Taylor, il nous faut les dérivées :

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

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

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

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

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.

En 2-D, la Gaussienne 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 séparable et symétrique circulaire.

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.

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

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

Les Filtres Numérique Gaussien

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 $N \leq 7$, les "ondes" de $W_R(f)$ dominent le spectre.
Pour $N \geq 9$, 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. $G_x(i, j, \sigma)$

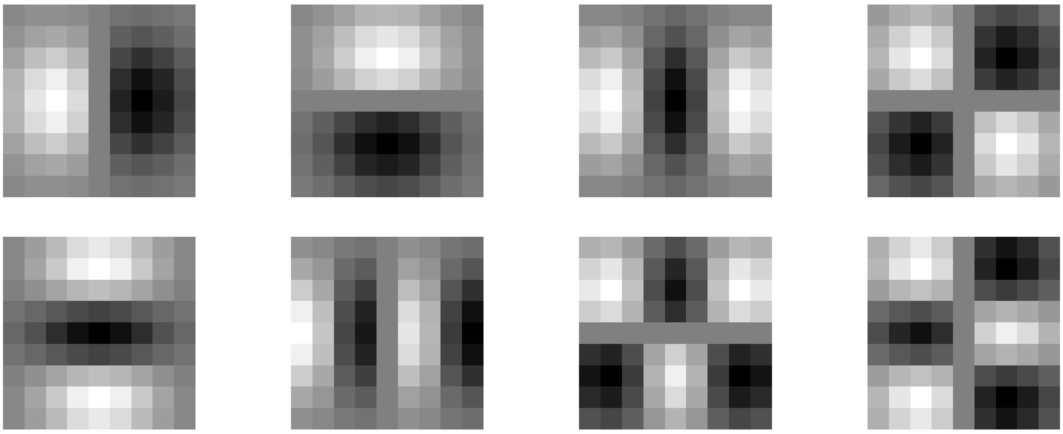
$$G_x(i, j, \sigma) = -\frac{i}{\sigma^2} \; G(i, j, \sigma) = -\frac{i}{\sigma^2} \; G(i, \sigma) * G(j, \sigma)$$

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

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(i, j), $\vec{\ }$ p(i,j) est calculé par $\vec{\ }$ $G(i, j, \sigma) * p(i,j)$.

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

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: $^2p(i,j) = G_{xx}(i, j, \sigma)*p(i, j) + G_{yy}(i, j, \sigma) * p(i, j)$
(La Laplacien est une scalaire.)