

AMO : Analyse Multiéchelles et Ondelettes

Dominique Béréziat

Master Informatique, spécialité IMA
Université Pierre et Marie Curie

16 janvier 2014

Espace d'échelles

Plan du cours

- ▶ 3 cours : espace d'échelles linéaires continus, cas discret, espaces d'échelles non linéaire.
- ▶ Comment faire en sorte que les opérateurs différentiels fonctionnent à **différentes échelles** ?
- ▶ Pourquoi doit-on régulariser certains problèmes ?
- ▶ Pourquoi régulariser équivaut à lisser ?
- ▶ Pourquoi le cas continu est très différent du cas discret (unicité notamment) ?
- ▶ Pourquoi les problèmes du traitement d'images formulés comme des EDP à résoudre (ou une fonction de coût à minimiser avec ou sans contrainte) engendrent des espaces d'échelles ?

Chapitre 1 : espaces d'échelles continus

Introduction

Espace d'échelles continus

Applications

Introduction

Motivations

Premiers travaux : décomposition d'images

Espace d'échelles continus

Applications

Motivations

- ▶ Dans beaucoup de problèmes images :

Axiome

La détection d'une primitive se fait à une échelle optimale.

Axiome ([Marr and Hildreth, 1980])

Une image est caractérisée par ses variations locales d'intensité à chaque échelle.

- ▶ Exemples :
 - ▶ Calcul du gradient : différence finie, convolution gaussienne, ... → problème de portée de l'opérateur.
 - ▶ Détection de contours : dépend du gradient (ou du noyau de Canny) : toujours un problème de portée ..
 - ▶ Flot optique : dépend du gradient.
 - ▶ Segmentation : information *a priori* (exemple : taille ou forme des régions), du gradient ? des contours ?
 - ▶ Reconnaissance de formes ...

Terminologie

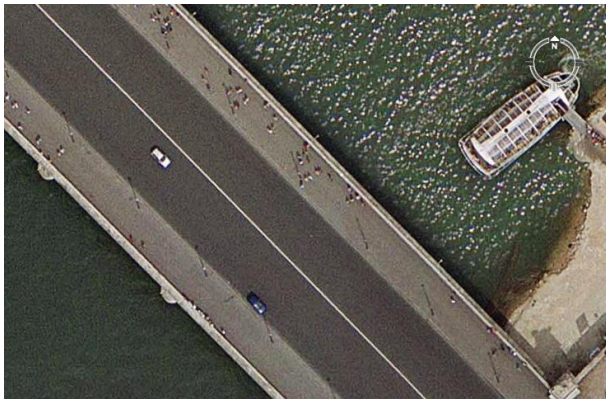
Des notions à préciser ...

- ▶ Une primitive (dans une image) :
 - ▶ des contours,
 - ▶ des structures de pixels,
 - ▶ des régions,
 - ▶ des attributs (simple à calculer : passage à zéro du laplacien),
 - ▶ ...
- ▶ L'échelle :
 - ▶ une primitive \Rightarrow une taille de représentation perceptible dans les images,
 - ▶ une mesure abstraite des structures existantes dans une image,
 - ▶ dépend fortement des modèles *a priori*,
 - ▶ exemple : télétection (Google Earth) ...

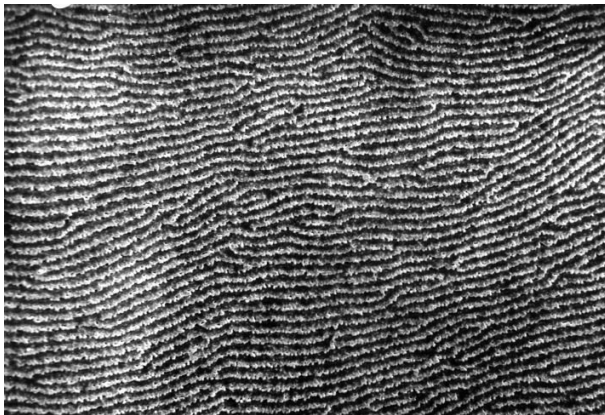
Exemple (télédetection ; Google Earth)



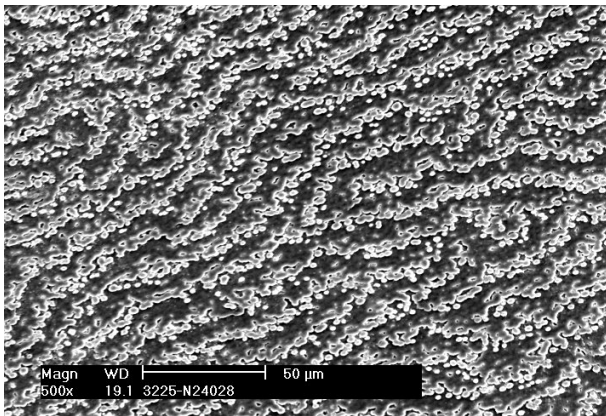
Exemple (télédetection ; Google Earth)



Exemple (imagerie électronique ; huîtres)

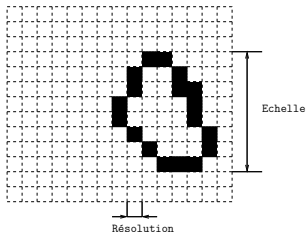


Exemple (imagerie électronique ; huîtres)



Terminologie (suite)

- ▶ La résolution :
 - ▶ le nombre de pixels dans une image,
 - ▶ donc échelle \neq résolution.



- ▶ Des algorithmes multi-échelles et/ou multi-résolutions.
- ▶ Parfois, on confond les deux notions !

Échelle / résolution

- ▶ Des structures apparaissent / disparaissent selon l'échelle de représentation spatiale des images (la résolution spatiale).
- ▶ Voir exemples Google Earth et huîtres.
- ▶ Dans les images : les intervalles d'échelles sont bornées :
 - ▶ une échelle minimale : conditionnée par la taille du pixel,
 - ▶ une échelle maximale : conditionnée par le nombre de pixel.

Notations, représentation discrète et continue

- ▶ Dans ce cours, on désignera par I une image, c'est-à-dire une fonction à deux inconnues, et par f un signal, c'est-à-dire une fonction à une inconnue.
- ▶ On utilisera tantôt la représentation discrète d'une image (ou d'un signal) :
 - ▶ Représentation discrète : I est définie sur un domaine discret et borné et donne ses valeurs également dans un domaine discret.
Typiquement :

$$I : \{1, \dots, N\} \times \{1, \dots, P\} \rightarrow \{0, \dots, L-1\}$$
$$(x, y) \mapsto I(x, y)$$

Ainsi, $x = 1, 2, \dots, N$ et $y = 1, \dots, P$.

- ▶ C'est la représentation naturelle et intuitive des images.

Notation (suite)

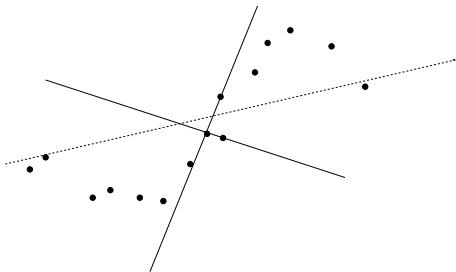
- ▶ Tantôt la représentation continue :
 - ▶ Représentation continue : la fonction I prends ses valeurs dans un domaine borné de \mathbb{R}^2 et donne ses valeurs dans \mathbb{R} . Typiquement

$$\begin{aligned} I : [0, 1] \times [0, 1] &\rightarrow [0, 1] \\ (x, y) &\mapsto I(x, y) \end{aligned}$$

- ▶ Le problème est souvent formulé en continue (I supposé dérivable/intégrable, utilisation des opérateurs différentiels et d'intégration)
- ▶ On notera souvent $\mathbf{x} = (x, y)$ pour alléger les équations.

Taille des opérateurs

- ▶ Passage du continu au discret : la discrétisation, on en parlera longuement dans le second cours.
- ▶ Le choix de la discrétisation des opérateurs est mal posé :



- ▶ Exemple de l'approximation de la dérivé :

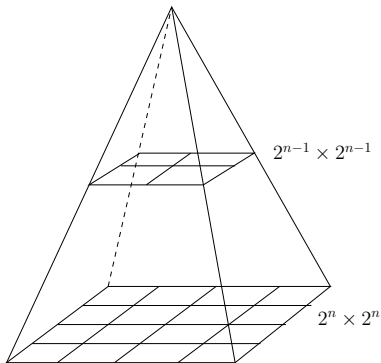
$$\begin{aligned}\frac{\partial}{\partial x} f(x_i) &\sim \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i} \\ &\sim \frac{f(x_{i+1}) - f(x_{i-1}))}{x_{i+1} - x_{i-1}}\end{aligned}$$

Première représentation des espaces d'échelles

- ▶ L'idée de représentation multi-échelles / multi-résolutions n'est pas neuve :
 - ▶ [Rosenfeld and Thurston, 1971] : utilisation d'opérateurs de différente taille pour la détection de contours.
 - ▶ [Klinger, 1971, Uhr, 1972] : représentation à différentes résolutions (obtenues par sous-échantillonnage) \Rightarrow représentation pyramidale

Représentation par pyramide

- ▶ Soit $K > 0$ entier, et $0 < k \leq K$ également entier.
- ▶ $I^K \equiv I$ image à résolution maximale.
- ▶ I^k image à niveau k est obtenue en :
 - ▶ sous-échantillonnant l'image I^{k-1}
 - ▶ moyennant le voisinage 2×2 de chaque pixel de I^{k-1}
- ▶ Chaque pixel de I^{K-i} correspond au moyennage de la zone $2^i \times 2^i$ de I^K .



Pyramides

- ▶ Inconvénient : le moyennage correspond à une opération de fort lissage.
 - ⇒ Grosses répercussions dans le domaine de Fourier.
- ▶ Amélioration : remplacer le moyennage par un lissage plus fin : gaussien par exemple.
- ▶ Opérateur de réduction \mathcal{R} (ici pour un signal) :

$$f^{(k-1)}(x) = \mathcal{R}(f^{(k)}) = \sum_{n \in \mathbb{N}} c(n) f^{(k)}(2x - n)$$

avec c filtre de lissage.

- ▶ Dans \mathbb{R}^2 (pour une image), on prendra un filtre séparable ($c(n, p) = c(n)c(p)$).

$$I^{(k-1)}(x, y) = \mathcal{R}(f^{(k)}) = \sum_{n \in \mathbb{N}} \sum_{p \in \mathbb{N}} c(n, p) I^{(k)}(2x - n, 2y - p)$$

Caractérisation de c (dans 1D)

► Propriétés spatiales :

- support fini dans \mathbb{Z} : $\exists N$ t.q. $\forall |n| > N, c(n) = 0$.
- positivité : $c(n) \geq 0$
- unimodalité : $c(|n|) \geq c(|n+1|)$
- symétrie : $c(n) = c(-n)$
- normalisation : $\sum c(n) = 1$
- contribution égale à chaque niveau $\sum c(2n) = \sum c(2n+1)$

► Propriétés fréquentielles :

- filtre passé-bas : élimination des hautes fréquences (donc des détails les plus petits).
- le filtre doit approximer un passe-bas idéal (la représentation fréquentielle d'un passe-bas n'est pas à support compact).

► Exemples :

- $N = 3$: filtre unique : $(\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$
- $N = 5$: $(\frac{1}{4} - \frac{a}{2}, \frac{1}{4}, a, \frac{1}{4}, \frac{1}{4} - \frac{a}{2})$ si $a = 0.4$ on est au plus proche d'une gaussienne.

Pyramides “passe-bande”

- ▶ Elles correspondent à un filtrage passe-bande (plutôt que passe-bas).
- ▶ Construction : différence entre deux niveaux adjacents d'une pyramide passe-bas :

$$L^{(k)} = f^{(k)} - \mathcal{E}(f^{(k-1)}) \quad (1)$$

$$L^{(0)} = f^{(0)} \quad (2)$$

- ▶ \mathcal{E} opération d'expansion, pour “revenir” à la résolution précédente == suréchantillonnage
- ▶ mais l'info complète du niveau précédent n'est évidemment pas retrouvée.

$$\mathcal{E}(f^{(k-1)}) = 2 \sum c(n) f^{(k-1)} \left(\frac{x-n}{2} \right)$$

Applications des décompositions pyramidales passe-bande

- ▶ Détection de structures : les pyramides passe-bande localisent les structures de taille compatible avec le niveau de décomposition.
- ▶ L'information n'est pas redondante pour chaque niveau (décorrélation entre chaque niveau) : chaque niveau peut être compressé efficacement.
- ▶ On peut garder certains niveaux de détail.
- ▶ Une reconstruction (imparfaite) est possible en inversant (1, 2) :

$$\begin{aligned}\tilde{f}^{(0)} &= L^{(0)} \\ \tilde{f}^{(k)} &= L^{(k)} + \mathcal{E}(\tilde{f}^{(k-1)})\end{aligned}$$

Avantages / Inconvénients

- ▶ Algos faciles à implémenter, de faible complexité.
- ▶ Permet de réduire les coûts (en diminuant la résolution donc) \Rightarrow algos incrémentaux.
- ▶ Inconvénients : décomposition non invariante par translation, par rotation (pb pour certaines applications).
- ▶ Pour en savoir plus : [Burt, 1981], [Crowley, 1981].

Introduction

Espace d'échelles continus

Première définition dans \mathbb{R}

Espace d'échelles pour \mathbb{R}^2

Propriété du noyau gaussien 1D

Équation de la chaleur

Dérivation dans les espaces d'échelles

Unicité de la représentation en espace d'échelles

Lien avec les ondelettes

Applications

Première définition (dans \mathbb{R})

- ▶ **Cas continu** : le paramètre d'échelle et le domaine spatial sont continus.
- ▶ Les espaces d'échelles ont été initiés à partir de l'article fondateur [Witkin, 1983].
- ▶ L'article traite des signaux 1D ; à chaque signal f , on associe une famille L de signaux indexée par le paramètre $t \in \mathbb{R}^+$:

$$f : \mathbb{R} \rightarrow \mathbb{R} \quad \text{signal original}$$

$$L : \mathbb{R} \times \mathbb{R}^+ \rightarrow \mathbb{R} \quad \text{famille qui dérive de } f$$

- ▶ Définition de L :

$$t = 0 \quad L(., t = 0) = f$$

$$t > 0 \quad L(., t) = g_{\sigma(t)} \star f$$

Choix de g

- ▶ g est un noyau de convolution (filtrage).
- ▶ Fonction gaussienne d'écart-type σ :

$$g_{\sigma} = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$$

- ▶ Remarque : $\sigma(t) = \sqrt{t}$.
- ▶ Effet sur un signal :

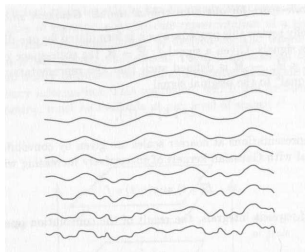


Figure : From [Witkin, 1983]

Une dimension supplémentaire

- ▶ C'est un lissage gaussien.
- ▶ On construit un espace de dimension 3 autour de l'image (espace de dimension 2).
- ▶ Dimension supérieure : facteur d'échelle t :

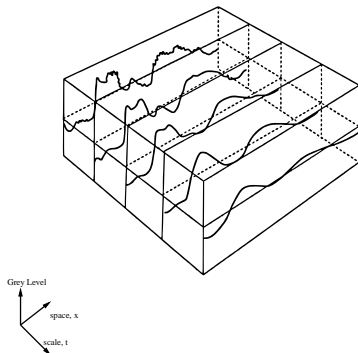


Figure : From [Witkin, 1983]

Lien avec l'équation de la chaleur

- Équation de la chaleur (L =température) :

$$\frac{\partial}{\partial t} L(x, t) = \frac{1}{2} \frac{\partial^2}{\partial x^2} L(x, t) \quad (3)$$

$$L(x, 0) = f(x) \quad (\text{condition initiale}) \quad (4)$$

- Cette équation régit la diffusion de la chaleur dans un milieu 1-D homogène.

Théorème

Une solution du système d'équation (3,4) est donnée par :

$$L(x, t) = g_{\sqrt{t}} \star f(x)$$

où g est la gaussienne d'écart-type \sqrt{t} .

Preuve

- On vérifie aisément que $g_{\sqrt{t}}$ est solution de (3) :

$$\begin{aligned}g(x, t) &= \frac{1}{t} e^{-\frac{x^2}{2t}} \\ \frac{\partial}{\partial t} g(x, t) &= -\frac{1}{2t^2} e^{-\frac{x^2}{2t}} + \frac{1}{t} \left(\frac{x^2}{2t^2} \right) e^{-\frac{x^2}{2t}} \\ &= \frac{1}{2t^2} e^{-\frac{x^2}{2t}} \left[\frac{x^2}{t} - 1 \right] \\ \frac{\partial}{\partial x} g(x, t) &= -\frac{x}{t^2} e^{-\frac{x^2}{2t}} \\ \frac{\partial^2}{\partial x^2} g(x, t) &= -\frac{1}{t^2} e^{-\frac{x^2}{2t}} - \frac{x}{t^2} \left(-\frac{x}{t} \right) e^{-\frac{x^2}{2t}} \\ &= \frac{1}{t^2} e^{-\frac{x^2}{2t}} \left[\frac{x^2}{t} - 1 \right] \\ \Rightarrow \frac{\partial}{\partial t} g(x, t) &= \frac{1}{2} \frac{\partial^2}{\partial x^2} g(x, t)\end{aligned}$$

Preuve (suite)

- ▶ $g_{\sqrt{t}}(x)$ n'est pas solution de (4) (indétermination en $t = 0$).

Lemme

$$\begin{aligned}\frac{\partial}{\partial x}(f \star g)(x) &= f \star \frac{\partial g}{\partial x}(x) \\ f \star (g + h) &= f \star g + f \star h\end{aligned}$$

- ▶ En remplaçant L par $f \star g_{\sqrt{t}}$ dans l'équation (3), et en appliquant le lemme, on remarque que $f \star g$ est solution de (3) :

$$\begin{aligned}\frac{\partial}{\partial t}(f \star g_{\sqrt{t}}) - \frac{1}{2} \frac{\partial^2}{\partial x^2}(f \star g_{\sqrt{t}}) &= f \star \frac{\partial}{\partial t} g_{\sqrt{t}} - \frac{1}{2} f \star \frac{\partial^2}{\partial x^2} g_{\sqrt{t}} \\ &= f \star \left(\frac{\partial}{\partial t} g_{\sqrt{t}} - \frac{1}{2} \frac{\partial^2}{\partial x^2} g_{\sqrt{t}} \right) \\ &= 0\end{aligned}$$

Preuve (suite)

- ▶ Reste à montrer que (4) est vérifiée.
- ▶ $g_{\sqrt{t}} \star f(x) = \int_{\mathbb{R}} f(x-u)g_{\sqrt{t}}(u)du$
- ▶ Résultat admis :

$$\lim_{t \rightarrow 0} \int_{\mathbb{R}} f(x) \frac{1}{t\sqrt{2\pi}} e^{-\frac{x^2}{2t}} dx = \int_{\mathbb{R}} f(x)\delta(x)dx$$

pour toute fonction f intégrable dans \mathbb{R} .

- ▶ La mesure de Dirac (δ) est une fonction “spéciale” définie indirectement par :

$$\int_{\mathbb{R}} f(x)\delta(x)dx = f(0)$$

- ▶ Donc, par définition de δ , on a :

$$\lim_{t \rightarrow 0} \int_{\mathbb{R}} f(x-u)g_{\sqrt{t}}(u)du = \int_{\mathbb{R}} f(x-u)\delta(u)du = f(x)$$

Preuve du Lemme

- ▶ distributif par rapport aux dérivées partielles :

$$\begin{aligned}\frac{\partial}{\partial x}(f \star g)(x) &= \frac{\partial}{\partial x} \int_{\mathbb{R}} f(y)g(x-y)dy \\ &= \int_{\mathbb{R}} \frac{\partial}{\partial x}(f(y)g(x-y))dy \\ &= \int_{\mathbb{R}} f(y) \frac{\partial}{\partial x}g(x-y)dy\end{aligned}$$

- ▶ linéarité de la convolution :

$$\begin{aligned}f \star (g + h) &= \int_{\mathbb{R}} f(x-y)(g(y) + h(y))dy \\ &= \int_{\mathbb{R}} f(x-y)g(y)dy + \int_{\mathbb{R}} f(x-y)h(y)dy\end{aligned}$$

Preuve du Lemme

- pour mémoire :

$$\begin{aligned} f \star g(x) &= \int_{\mathbb{R}} f(x-y)g(y)dy \\ &= \int_{\mathbb{R}} f(y)g(x-y)dy \end{aligned}$$

par changement de variable (inversion du domaine). Donc le produit de convolution est commutatif.

Conclusion

- ▶ Le filtrage gaussien lisse les détails de plus en plus fortement avec t croissant,
- ▶ Question : peut-on prendre n'importe quel filtre ?
- ▶ Une propriété recherchée : si on augmente le paramètre d'échelle, on ne doit pas voir apparaître de “nouvelles structures” (propriété de causalité).
- ▶ En particulier, le bruit ne doit pas faire apparaître des structures non significantes.
- ▶ La propriété de causalité est la propriété centrale des espaces d'échelles et a été formalisée par [Koenderink, 1984].
- ▶ Généralisation à \mathbb{R}^2 ?

Espace d'échelles pour \mathbb{R}^2

Définition (Espace d'échelles pour une image continue)

Soit $I : \mathbb{R}^2 \rightarrow \mathbb{R}$ une image. On appelle représentation d'une image I en espace d'échelles linéaire la fonction L telle que :

$$\begin{aligned}L &: \mathbb{R}^2 \times \mathbb{R}^+ \rightarrow \mathbb{R} \\L(\mathbf{x}, 0) &= I(\mathbf{x}), \forall \mathbf{x} \in \mathbb{R}^2 \\L(\mathbf{x}, t) &= g_{\sqrt{t}} \star I(\mathbf{x}), \forall \mathbf{x} \in \mathbb{R}^2, t \in \mathbb{R}^+\end{aligned} \quad (5)$$

avec :

$$g_{\sqrt{t}}(\mathbf{x}) = \frac{1}{2\pi t} e^{-\frac{x^2+y^2}{2t}}$$
$$f \star g(\mathbf{x}) \equiv \int_{\mathbb{R}^2} g(x, y) f(x' - x, y' - y) dx' dy'$$

Exemple



(a) Image originale



(b) $t = 1.0$



(c) $t = 1.5$

Figure : Exemple de décomposition

- Disparition des détails au fur et à mesure du lissage ...

Exemple (animation)

Loading data ...

Figure : La même chose en séquence ...

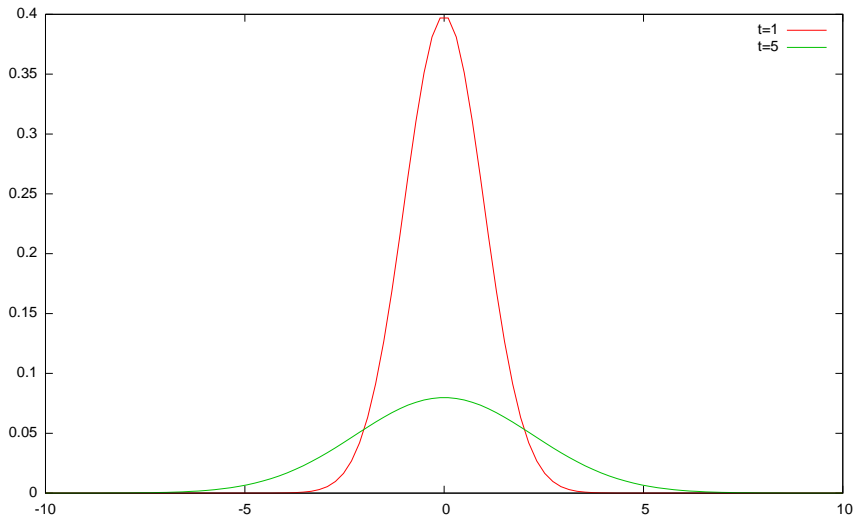
Étude du noyau gaussien

- ▶ Le noyau gaussien lisse : il ne crée pas de nouvelles structures.
- ▶ Filtre moyennneur pondéré :

$$\int_{\mathbb{R}} g_{\sigma}(x) dx = 1 \quad (6)$$

- ▶ Décroissance rapide de g
 - ▶ Au delà de $|x| > \sqrt{t}$, $g(x) \sim 0$.
- ⇒ Les détails plus petits que \sqrt{t} sont supprimés.

Grphe de la gaussienne



Preuve équation (6)

- Calcul de $\int e^{-x^2} dx$:

$$\begin{aligned}\left(\int_{\mathbb{R}} e^{-x^2} dx\right)^2 &= \left(\int_{\mathbb{R}} e^{-x^2} dx\right) \left(\int_{\mathbb{R}} e^{-y^2} dy\right) \\ &= \iint_{\mathbb{R}^2} e^{-x^2-y^2} dx dy\end{aligned}$$

- Changement de variable (polaire) : $x = r \cos \theta$ et $y = r \sin \theta$

$$\begin{aligned}\left(\int_{\mathbb{R}} e^{-x^2} dx\right)^2 &= \int_{\mathbb{R}^+} \int_0^{2\pi} e^{-r^2} \begin{vmatrix} \cos \theta & \sin \theta \\ -r \sin \theta & r \cos \theta \end{vmatrix} dr d\theta \\ &= \int_0^{2\pi} \int_0^{+\infty} re^{-r^2} dr d\theta = \int_0^{2\pi} d\theta \int_0^{+\infty} re^{-r^2} dr \\ &= 2\pi \left[-\frac{1}{2} e^{-r^2} \right]_{r=0}^{r=+\infty} = \pi\end{aligned}$$

Une structure de semi-groupe

- ▶ La famille des fonctions gaussiennes $(F, g \star)$ forme un semi-groupe.
 - ▶ On a la relation : $g(\cdot, t) \star g(\cdot, s) = g(\cdot, t + s)$
- ⇒ Une représentation à l'échelle t_2 peut être déduite de n'importe quelle représentation à une échelle t_1 inférieur :

$$t_2 > t_1, L(t_2, x) = g_{t_2-t_1} \star L(t_1, x)$$

- ▶ Dans l'espace de Fourier, on utilise la relation (Théorème de convolution/Fourier) :

$$\widehat{L}_{t_2}(w) = \widehat{g}_{t_2-t_1}(w) \times \widehat{L}_{t_1}(w)$$

Fonction gaussienne dans \mathbb{R}^2

- ▶ L'extension est naturelle car le filtre est séparable :

$$\begin{aligned}g_{\sigma}(\mathbf{x}) &= g_{\sigma}(x) \times g_{\sigma}(y) \\ &= \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t} \frac{1}{\sqrt{2\pi t}} e^{-y^2/2t} = \frac{1}{2\pi t} e^{-(x^2+y^2)/2t}\end{aligned}$$

- ▶ L'implémentation d'un filtre séparable est facile à mettre en œuvre car :

$$I \star g_{\sigma}(\mathbf{x}) = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} I(x - x', y - y') g_{\sigma}(x') dx' \right) g(y') dy'$$

- ▶ Cette propriété permet d'étendre facilement l'équation (3) à \mathbb{R}^n ...

Équation de la chaleur

Théorème

La représentation en espace d'échelles L d'une fonction f de \mathbb{R}^2 vérifie le système d'équations suivant :

$$\frac{\partial}{\partial t} L(\mathbf{x}, t) = \frac{1}{2} \nabla^2 L(\mathbf{x}, t) \quad (7)$$

$$L(\mathbf{x}, 0) = f(\mathbf{x}) \quad (8)$$

- ▶ On note $\nabla^2 L = \sum_{i=1}^2 \frac{\partial^2}{\partial x_i^2} L$, l'opérateur laplacien.
- ▶ Le système (7,8) correspond à l'équation de la chaleur dans \mathbb{R}^2 pour un milieu homogène.

Principe du maximum (illustré en 1D)

- ▶ Le principe fondamental de non création de structures avec l'échelle croissante est vérifié.
- ▶ On part de l'hypothèse qu'une structure est caractérisée par ses contours, donc par des extrema dans l'image.
 - ▶ à l'échelle t_0 , si x_0 est un maximum local, alors
$$\nabla^2 L(x_0, t_0) < 0 \Rightarrow \frac{\partial}{\partial t} L(x_0, t_0) < 0$$
 - ▶ D.L. en t : $L(x_0, t) \sim L(x_0, t_0) + \frac{\partial}{\partial t} L(x_0, t_0)(t - t_0)$
 - ▶ Donc : pour $t > t_0$, $L(x_0, t) < L(x_0, t_0)$
 - ▶ à l'échelle t_0 , si x_0 minimum local, alors
$$\nabla^2 L(x_0, t_0) > 0 \Rightarrow \frac{\partial}{\partial t} L(x_0, t_0) > 0$$
 - ▶ DONC pour $t > t_0$, $L(x_0, t) > L(x_0, t_0)$
- ▶ Les maxima décroissent, les minima croissent : homogénéisation de l'image.

Dérivation dans les espaces d'échelles

- Considérons une image à l'échelle t , calculons sa différentielle :

$$\begin{aligned}\frac{\partial}{\partial x} L(x, t) &= \frac{\partial}{\partial x} (I \star g_{\sqrt{t}}(x)) \\ &= L \star \frac{\partial}{\partial x} g_{\sqrt{t}}(x)\end{aligned}$$

- La fonction $g_{\sqrt{t}}$ est C^∞ et $\frac{\partial^n}{\partial x^n} g_{\sqrt{t}}$ est intégrable $\forall n$.
- Il suffit que I soit intégrable (mais pas nécessairement dérivable) pour que $\frac{\partial^n}{\partial x^n} L$ soit défini.
- Cela donne une définition *faible* de la dérivée (on peut "dériver" une fonction non dérivable) :

$$\begin{aligned}I \star g_\sigma &\xrightarrow{\sigma \rightarrow 0} I \\ I \star \frac{\partial^n}{\partial x^n} g_\sigma &\xrightarrow{\sigma \rightarrow 0} \frac{\partial^n}{\partial x^n} I\end{aligned}$$

Propriété des dérivées dans l'espace d'échelles

- ▶ Les propriétés des espaces d'échelles s'appliquent aux images dérivées.
- ▶ En effet : si \mathcal{L} est solution de $\frac{\partial}{\partial t}L = \frac{1}{2}\nabla^2L$ alors $\partial\mathcal{L}$ est aussi une solution.
- ▶ C'est une conséquence immédiate de la commutation de la convolution avec les opérateurs différentiels.
- ▶ De plus, on a :

$$\mathcal{G}_{\sqrt{t_1}} \star \frac{\partial^n}{\partial X^n} \mathcal{G}_{\sqrt{t_2}} = \frac{\partial^n}{\partial X^n} \mathcal{G}_{\sqrt{t_1+t_2}} \quad (9)$$

$$\frac{\partial^m}{\partial X^m} \mathcal{G}_{\sqrt{t_1}} \star \frac{\partial^n}{\partial X^n} \mathcal{G}_{\sqrt{t_2}} = \frac{\partial^{n+m}}{\partial X^{n+m}} \mathcal{G}_{\sqrt{t_1+t_2}} \quad (10)$$

Dérivation de g : formule d'Hermitte

- ▶ On peut calculer explicitement la dérivée d'ordre n de la gaussienne à l'aide de la formule d'Hermitte :

$$\frac{\partial^n}{\partial x^n} e^{-x^2} = (-1)^n H_n(x) e^{-x^2} \quad (11)$$

- ▶ H_n est le polynôme d'Hermitte calculé par la formule récurrente :

$$\begin{cases} H_0(x) = 1 \\ H_{n+1}(x) = xH_n(x) - H'_n(x) \end{cases} \quad (12)$$

(On a : $H_1(x) = x$, $H_2(x) = x^2 - 1$, $H_3(x) = x^3 - 3x$, ...)

- ▶ *In fine* :

$$\frac{\partial^n}{\partial x^n} g_{\sqrt{t}}(x) = (-1)^n \frac{1}{(2t)^{\frac{n}{2}}} H\left(\frac{x}{\sqrt{2t}}\right) g_{\sqrt{t}}(x) \quad (13)$$

Unicité de la représentation en espace d'échelles [Lindeberg, 1994]

- ▶ Nous donnons ici une preuve d'unicité du noyau gaussien comme représentation des espaces d'échelles sous les hypothèses suivantes :
 - ▶ La solution est recherchée sous la forme d'un noyau de convolution : ainsi la linéarité est vérifiée et la représentation est invariante par translation (faire un changement de variable dans l'intégrale de la convolution).
 - ▶ La famille de solution forme un semi-groupe.

Théorème (Unicité du noyau gaussien)

Les solutions de l'équation (7, 8) qui s'écrivent sous la forme

$$L(\mathbf{x}, t) = h_t \star l(\mathbf{x}) = \int_{\Omega} h_t(\mathbf{x}') f(\mathbf{x} - \mathbf{x}') d\mathbf{x}' \quad (14)$$

ont comme noyau de convolution h_t une fonction gaussienne en \mathbf{x} de variance t .

Unicité du noyau gaussien (suite)

- ▶ La preuve repose sur le principe physique suivant (admit) :

Lemme (Pi-théorème (Vaschey-Buckingham))

Soit un système physique reliant n variables d'états (avec dimension) tel que $f(q_1, \dots, q_n) = 0$. Alors il existe une famille de paramètres π_i sans dimension et une fonction F telles que :

$$\begin{cases} F(\pi_1, \dots, \pi_p) = 0 \\ \pi_i = \prod_{k=1}^n q_k^{l_k} \end{cases} \quad (15)$$

Les variables π_i sont sans dimension car elle s'expriment comme des produits des variables q_j .

- ▶ Ce lemme signifie qu'un principe physique ne dépend jamais du choix des unités (et donc de l'échelle).

Unicité du noyau gaussien (preuve)

- ▶ On pose $L^{(t)}(\mathbf{x}) = L(\mathbf{x}, t)$.
- ▶ Dans l'espace de Fourier, l'équation (14) s'écrit :

$$\begin{aligned}\widehat{L}^{(t)}(\mathbf{w}) &= \widehat{h_t} \star \widehat{l}(\mathbf{w}) = \widehat{h_t}(\mathbf{w})\widehat{l}(\mathbf{w}) \\ \frac{\widehat{L}^{(t)}(\mathbf{w})}{\widehat{l}(\mathbf{w})} &= \widehat{h_t}(\mathbf{w})\end{aligned}$$

avec $\mathbf{w} = (u, v)$ coordonnées dans l'espace de Fourier.

- ▶ Le paramètre \mathbf{w} est une fréquence, donc \mathbf{w}^{-1} est une période.
- ▶ Le paramètre \sqrt{t} mesure une échelle.
- ▶ Ce sont deux longueurs donc $\mathbf{w}\sqrt{t}$ est sans dimension.
- ▶ Idem pour $\frac{\widehat{L}^{(t)}(\mathbf{w})}{\widehat{l}(\mathbf{w})}$.
- ▶ Lemme \Rightarrow on peut écrire que $\widehat{h_t}(\mathbf{w}) = \widehat{H}(\sqrt{t}\mathbf{w})$.

Unicité du noyau gaussien (preuve)

- ▶ On travaille avec : $\frac{\widehat{L}^{(t)}(\mathbf{w})}{\widehat{l}(\mathbf{w})} = \widehat{H}(\mathbf{w}\sqrt{t})$.
- ▶ $I = L^{(0)}$ donc $\widehat{H}(0) = 1$.
- ▶ Semi-groupe : $\widehat{h}_{t_1}(\mathbf{w})\widehat{h}_{t_2}(\mathbf{w}) = \widehat{h}_{t_1+t_2}(\mathbf{w})$ donc :

$$\widehat{H}(\mathbf{w}\sqrt{t_1})\widehat{H}(\mathbf{w}\sqrt{t_2}) = \widehat{H}(\mathbf{w}\sqrt{t_1+t_2})$$

- ▶ Posons : $\check{H}(\mathbf{w}^T \mathbf{w}) = \widehat{H}(\mathbf{w})$, $\mathbf{w} \in \mathbb{R}^2$ donc :

$$\begin{aligned} \check{H}((\mathbf{w}\sqrt{t_1})^T (\mathbf{w}\sqrt{t_1})) &\times \\ \check{H}((\mathbf{w}\sqrt{t_2})^T (\mathbf{w}\sqrt{t_2})) &= \check{H}((\mathbf{w}\sqrt{t_1+t_2})^T (\mathbf{w}\sqrt{t_1+t_2})) \\ \check{H}(t_1 \mathbf{w}^T \mathbf{w})\check{H}(t_2 \mathbf{w}^T \mathbf{w}) &= \check{H}((t_1+t_2)\mathbf{w}^T \mathbf{w}) \\ \check{H}(v_1)\check{H}(v_2) &= \check{H}(v_1+v_2) \end{aligned}$$

(en posant $v_1 = t_1 \mathbf{w}^T \mathbf{w}$ et $v_2 = t_2 \mathbf{w}^T \mathbf{w}$)

Unicité du noyau gaussien (preuve)

- Bilan :

$$\begin{cases} \tilde{H}(0) = 1 \\ \tilde{H}(v_1)\tilde{H}(v_2) = \tilde{H}(v_1 + v_2) \end{cases}$$

- C'est l'une des définitions de l'exponentielle ($u \mapsto e^{\alpha u}$).
- Donc $\hat{H}(\mathbf{w}) = e^{\alpha t \mathbf{w}^T \mathbf{w}}$.
- On prend $\alpha < 0$ ainsi $\lim_{\infty} \hat{H} = 0$.
- La transformée de Fourier d'une gaussienne est une gaussienne !
Plus précisément ($\alpha > 0$) :

$$\int_{\mathbb{R}^2} e^{-\frac{\mathbf{x}^T \mathbf{x}}{\alpha t}} e^{-i \mathbf{w} \mathbf{x}} d\mathbf{x} = e^{-\alpha t \mathbf{w}^T \mathbf{w}}$$

- On a des conditions de normalisation $\int_{\mathbb{R}} e^{-x^2/2t} dx = \sqrt{2\pi t}$ et finalement, on prend $\alpha = \frac{1}{2}$

Lien avec les ondelettes

- ▶ Un signal peut être représenté de plusieurs façons. Nous en connaissons au moins deux :
 - ▶ Représentation spatiale :

$$f(t) = \int_{\mathbb{R}} f(u) \underbrace{\delta(t-u)}_{\text{base}} du$$

Cette représentation localise en espace.

- ▶ Représentation fréquentielle (Fourier) :

$$\hat{f}(t) = \int_{\mathbb{R}} f(\omega) \underbrace{e^{i\omega t}}_{\text{base}} d\omega$$

Cette représentation localise en fréquence.

Lien avec les ondelettes

- ▶ Mais il existe des décompositions intermédiaires qui localisent à la fois en espace et en fréquence : ce sont les ondelettes.
- ▶ Pour toute fonction f de $L^2(\Omega)$ et de norme 1 ($\int |f|^2 = 1$). On défini :
 - ▶ $c(f) = \int t|f(t)|^2 dt$: “le centre”,
 - ▶ $\Delta(f) = \int (t - c(f))^2 |f(t)|^2 dt$: “la largeur”.toujours défini pour f L^2 -intégrable.
- ▶ Un théorème fondamental (admis) :

Théorème (Inégalité de Heisenberg)

$$\Delta(f)\Delta(\hat{f}) \geq \frac{1}{2} \quad (16)$$

Lien avec les ondelettes

- ▶ La conséquence de ce théorème : on ne peut pas localiser aussi précisément que l'on souhaite un signal à la fois en espace et en fréquence.

Proposition

Les fonctions f qui atteignent la borne d'Heisenberg sont :

$$f_{t_0, \Delta t, \omega_0}(t) = A e^{-(t-t_0)^2/2\Delta t^2} e^{i\omega_0 t}$$

où A est une constante normalisatrice : $\int |f_{t_0, \Delta t, \omega_0}|^2 = 1$.

- ▶ Ces fonctions forment une base d'ondelettes appelée Ondelettes de Gabor.

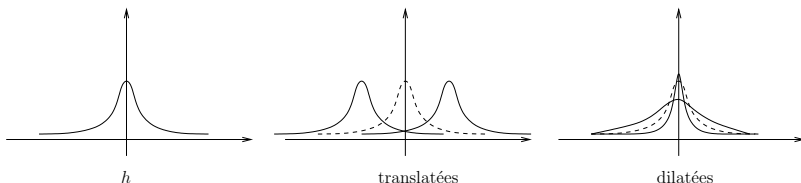
Définition des ondelettes

- ▶ Rappel : définition d'une ondelette :
- ▶ Une famille de fonctions $h_{a,b}$ telle que :

$$h_{a,b}(x) = \frac{h\left(\frac{x-b}{a}\right)}{\sqrt{|a|}}$$

$a, b \in \mathbb{R}, a \neq 0$

h est l'ondelette mère et la famille $h_{a,b}$ représente toutes les fonctions translatées et dilatées de h .



Définition des ondelettes (suite)

- ▶ Les ondelettes doivent vérifier la condition (dite d'admissibilité) sur h :

$$\int_{\mathbb{R}} \frac{|\hat{h}(\omega)|}{|\omega|} d\omega \leq \infty \quad (17)$$

signifie la décroissance rapide de la T.F.

- ▶ Soit f une fonction intégrable. On appelle transformée en ondelette continue la quantité :

$$\begin{aligned} (\mathcal{W}f)(a, b) &= \langle f, h_{a,b} \rangle = \int f(x) h_{a,b}(x) dx \\ &= \int f(x) h\left(\frac{x-b}{a}\right) \frac{dx}{\sqrt{a}} \end{aligned}$$

- ▶ La transformée en ondelettes est redondante : c'est une famille paramétrée par deux paramètres (un seul paramètre suffit pour représenter un signal).

Donc ... le lien ondelette/espace d'échelles

- ▶ On a $g_{0,0,1}(x) = e^{-\frac{x^2}{2}} = h(x)$: on limite la redondance à l'échelle.
- ▶ Et donc $h_{a,b}(x) = \frac{e^{-\frac{(x-b)^2}{2a}}}{\sqrt{a}}$.
- ▶ Donc la transformée en ondelette :

$$(\mathcal{W}f)(a, b) = \int f(x) \frac{e^{-\frac{(x-b)^2}{2a}}}{\sqrt{a}} dx$$

c'est bien une représentation en espace d'échelles !

- ▶ Enfin, la dérivée seconde de la gaussienne est l'ondelette mère de Marr.

Introduction

Espace d'échelles continus

Applications

- Implantation de la convolution gaussienne

- Détection de contours

- Détecteur de Harris multiéchelle

- Détection de *blobs*

- Détection de motifs invariant par changement d'échelle

- Sélection automatique de l'échelle

Implantation de la convolution Gaussienne

Matlab

```
function g = gD( f , scale , ox , oy)
%%% Perform a gaussian derivative convolution
%%% f: input image
%%% scale: smoothing parameter
%%% ox,oy: x and y derivate order (0,1,2,3)

% Build the gaussian kernel
K = ceil(3*scale);
x = -K:K;
Gs = exp(-x.^2/(2*scale ^2));
Gs = Gs/sum(Gs);

% Calculate the derivatives en x and y-direction
Gsx = gDerivative(ox,x,Gs, scale);
Gsy = gDerivative(oy,x,Gs, scale);

% Do the convolution
g = convSepBrd( f , Gsx , Gsy);
```

Implantation

Matlab

```
function r = gDerivative( order , x, Gs, scale)
%%% Compute a derivative of a 1D Gaussian kernel
%%% order: order derivation 0,1,2,3
%%% Gs: discretized , centered gaussian kernel
%%% scale: variance of Gs
switch order
  case 0
    r = Gs;
  case 1
    r = -x/(scale^2) .* Gs;
  case 2
    r = (x.^2-scale^2)/(scale^4) .* Gs;
  case 3
    r = x.*(3*scale^2-x.^2)/(scale^6) .* Gs;
  otherwise
    error('only derivatives up to third order are supported');
end
```

Implantation

Matlab

```
function g = convSepBrd( f, w1, w2)  
%%% convolve along cols and rows with repetition  
%%% of the border
```

```
N = size(f,1);  
M = size(f,2);  
K = (size(w1(:),1) - 1)/2;  
L = (size(w2(:),1) - 1)/2;  
iind = min(max((1:(N+2*K)) - K, 1), N);  
jind = min(max((1:(M+2*L)) - L, 1), M);  
fwb = f(iind, jind);  
g = conv2(w1, w2, fwb, 'valid');
```

Implantation



► `phone=imread("phone.pgm"); imshow(phone)`

Implantation



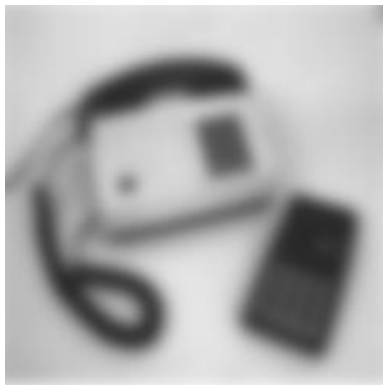
► `imshow(gD(phone, 1, 0, 0));`

Implantation



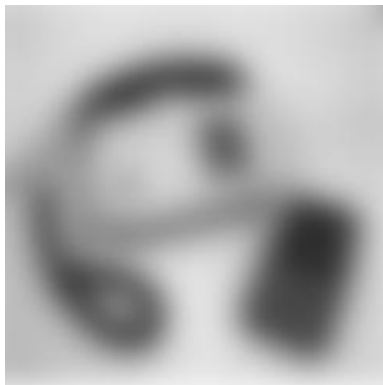
► `imshow(gD(phone,2,0,0));`

Implantation



► `imshow(gD(phone,5,0,0));`

Implantation



► `imshow(gD(phone,10,0,0));`

Implantation



► `imshow(gD(phone,1,1,0));`

Implantation



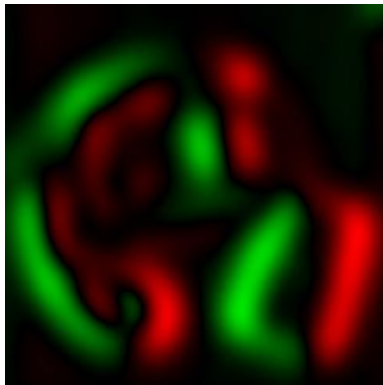
► `imshow(gD(phone,2,1,0));`

Implantation



► `imshow(gD(phone,5,1,0));`

Implantation



► `imshow(gD(phone,10,1,0));`

Implantation



► `imshow(gD(phone,1,0,1));`

Implantation



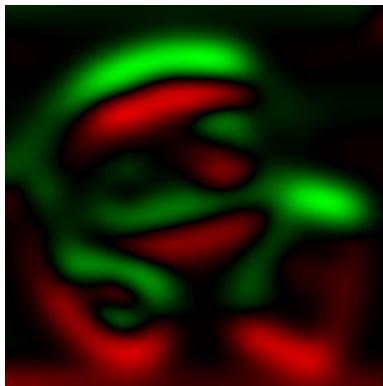
► `imshow(gD(phone,2,0,1));`

Implantation



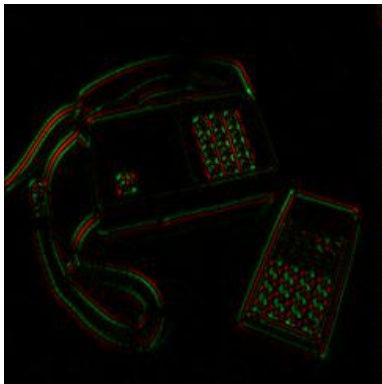
► `imshow(gD(phone,5,0,1));`

Implantation



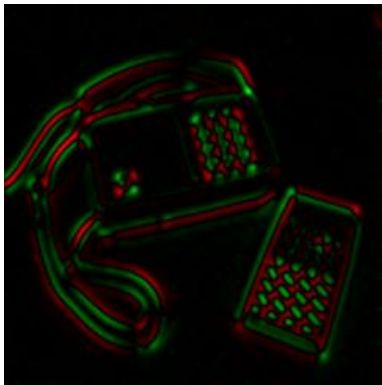
► `imshow(gD(phone,10,0,1));`

Implantation



► `imshow(gD(phone,1,1,1));`

Implantation



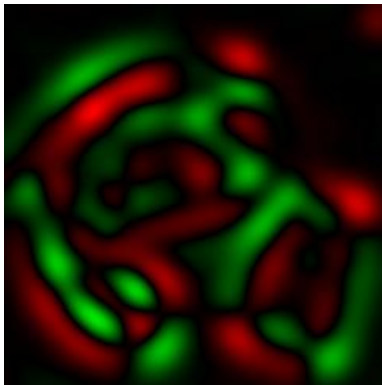
► `imshow(gD(phone,2,1,1));`

Implantation



► `imshow(gD(phone,5,1,1));`

Implantation



► `imshow(gD(phone,10,1,1));`

Application : détection de contours

- ▶ Une méthode de calcul de contours (alternative à Sobel/Canny/...)
- ▶ On utilise un système de coordonnées locales dépendant du gradient d'une image L :

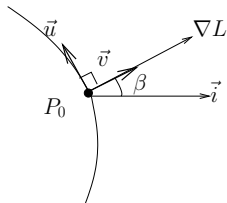


Figure : Coordonnées directionnelles

- ▶ Notons $L_x = \frac{\partial L}{\partial x}$, ...
- ▶ $\vec{v} = \frac{\nabla L}{\|\nabla L\|} = \frac{1}{\sqrt{L_x^2 + L_y^2}} \begin{pmatrix} L_x \\ L_y \end{pmatrix}$, $\vec{u} = \frac{\nabla^\perp L}{\|\nabla L\|} = \frac{1}{\sqrt{L_x^2 + L_y^2}} \begin{pmatrix} -L_y \\ L_x \end{pmatrix}$
- ▶ $\beta = \widehat{(\vec{v}, \vec{i})}$

Coordonnées directionnelles

⇒ Expression de \vec{u} et \vec{v} dans la base canonique.

► $\vec{v} = \begin{pmatrix} \cos \beta \\ \sin \beta \end{pmatrix}$ et $\begin{cases} \vec{u} = -\vec{i} \sin \beta + \vec{j} \cos \beta \\ \vec{v} = \vec{i} \cos \beta + \vec{j} \sin \beta \end{cases}$

► On peut définir maintenant :

$$\begin{cases} \frac{\partial}{\partial \vec{u}} = \sin \beta \frac{\partial}{\partial x} - \cos \beta \frac{\partial}{\partial y} \\ \frac{\partial}{\partial \vec{v}} = \cos \beta \frac{\partial}{\partial x} + \sin \beta \frac{\partial}{\partial y} \end{cases}$$

► Ainsi : $\cos \beta = \frac{L_x}{\|\nabla L\|}$ et $\sin \beta = \frac{L_y}{\|\nabla L\|}$.

► Propriété 1 :

$$L_{\vec{u}} = 0 \tag{18}$$

□ En effet : $\frac{\partial L}{\partial \vec{u}} = \sin \beta \frac{\partial L}{\partial x} - \cos \beta \frac{\partial L}{\partial y} = 0$ par définition de β . ■

Propriétés des coordonnées directionnelles

- Propriété 2 : Dans ce repère, la courbure des isocourbes de niveau (L) s'exprime naturellement par

$$\kappa = \frac{L_{\bar{u}\bar{u}}}{L_{\bar{v}}} \quad (19)$$

- En effet, la courbure d'une courbe implicite d'équation $L(x, y) = c$ est donnée par

$$\kappa = \frac{L_x^2 L_{yy} + L_y^2 L_{xx} - 2L_x L_y L_{xy}}{(L_x^2 + L_y^2)^{\frac{3}{2}}}$$

- il suffit de développer les opérateurs $\partial_{\bar{u}}$ et $\partial_{\bar{v}}$ dans l'équation (19) pour vérifier qu'elles coïncident. ■

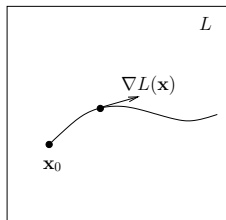
Propriétés des coordonnées directionnelles

- Propriété 3 : Dans ce repère, la courbure le long des lignes du gradient de L s'exprime naturellement par

$$\mu = \frac{L_{\bar{u}\bar{v}}}{L_{\bar{v}\bar{v}}} \quad (20)$$

- Les lignes de gradient sont les courbes que l'on peut tracer à partir d'un point en suivant le vecteur gradient. Elles sont définies par :

$$\begin{aligned} \mathbf{x}(0) &= \mathbf{x}_0 \\ \frac{\partial \mathbf{x}}{\partial t}(t) &= \nabla L(\mathbf{x}(t)) \end{aligned}$$



Détection de contours

- ▶ On utilise la définition de [Canny, 1986] pour caractériser un contour : maximum local de la norme du gradient dans la direction du gradient.
- ▶ Il faut donc examiner la magnitude de $L_{\bar{v}}$ (scalaire).
- ▶ $L_{\bar{v}}$ est un maximum local si :

$$\begin{aligned}L_{\bar{v}\bar{v}} &= 0 \quad (\text{extrema}) \\L_{\bar{v}\bar{v}\bar{v}} &< 0 \quad (\text{maxima})\end{aligned}$$

- ▶ En coordonnées cartésiennes :

$$\begin{aligned}L_{\bar{v}\bar{v}} &= \frac{L_x^2 L_{xx} + 2L_x L_y L_{xy} + L_y^2 L_{yy}}{L_x^2 + L_y^2} \\L_{\bar{v}\bar{v}\bar{v}} &= \frac{L_x^3 L_{xxx} + 3L_x^2 L_y L_{xxy} + 3L_x L_y^2 L_{xyy} + L_y^3 L_{yyy}}{L_x^2 + L_y^2}\end{aligned}$$

Calcul des gradients directionnels

- ▶ Calcul dans l'espace d'échelles.
- ▶ Par exemple :

$$\frac{\partial}{\partial \bar{u}} = \left(\frac{L_y^s}{\nabla L^s} \frac{\partial g_s}{\partial x} + \frac{L_x^s}{\nabla L^s} \frac{\partial g_s}{\partial y} \right) \star$$

$$\frac{\partial}{\partial \bar{v}} = \left(\frac{L_x^s}{\nabla L^s} \frac{\partial g_s}{\partial x} - \frac{L_y^s}{\nabla L^s} \frac{\partial g_s}{\partial y} \right) \star$$

avec $L^s = g_s \star L$

Résultat



Figure : Exemple avec $\sigma = 1$

Détection de contours (fin)

- ▶ Opérateur de Marr : passage par zéro du Laplacien.

$$\nabla^2 L = L_{xx} + L_{yy} = L_{\bar{u}\bar{u}} + L_{\bar{v}\bar{v}}$$

- ▶ En utilisant la courbure, on obtient :

$$\nabla^2 L = \kappa L_{\bar{v}} + L_{\bar{v}\bar{v}}$$

- ▶ L'opérateur de Marr permet de détecter des régions (et non pas des contours) qui correspondent à des extrema locaux de la norme du gradient.
- ▶ On peut *in fine* détecter des régions de différentes tailles selon l'échelle.

Détecteur de Harris

Rappel

- ▶ Le tenseur sur le gradient d'une image I .

$$A(I)(x, y) = G_\sigma \star \nabla I^T(x, y) \nabla I(x, y)$$

où G_σ est un noyau 2D gaussien de variance σ^2

- ▶ En pratique on considère plutôt

$$A(I) = \begin{pmatrix} G_\sigma \star (I_x^2) & G_\sigma \star (I_x I_y) \\ G_\sigma \star (I_x I_y) & G_\sigma \star (I_y^2) \end{pmatrix}$$

car sinon le déterminant de $\nabla I^T \nabla I$ est nul

- ▶ Soit $\kappa \in [0.04, 0.15]$, on définit alors l'opérateur de Harris par

$$R(I) = \det(A(I)) - \kappa \operatorname{Tr}^2(A(I))$$

- ▶ (x, y) est un coin $\Leftrightarrow R(I)(x, y) > 0$ et (x, y) est un maxima local de R .

Détecteur de Harris-Laplace

- ▶ Le détecteur de Harris est invariant par rotation :

$$R(I) = \lambda_1 \lambda_2 - \kappa(\lambda_1 + \lambda_2)^2$$

avec λ_1 et λ_2 les deux valeurs propres de $A(I)$.

- la convolution par G_σ rends $A(I)$ inversible (dès que σ est suffisamment grand).
- ▶ Mais R n'est pas invariant par changement d'échelle.
- ▶ Remplaçons $I(x, y)$ par $L(x, y, s) = G_s \star I(x, y)$ dans le calcul de A .
- ▶ On obtient le détecteur de Harris-Laplace. Le choix de l'échelle s en accord avec la taille du coin à localiser permet sa détection optimale par $R_{\kappa, \sigma, s}$.
- ▶ En pratique, $L_x(x, y, s)$ est calculé par $\frac{\partial G_s}{\partial x} \star I(x, y)$,

Détecteur de Harris-Laplace

Sélection automatique de l'échelle

► Algorithme itératif [Lindeberg, 1998] :

1. Soit $k = 1$, s_1 une valeur d'échelle initiale.
2. Soit (x_i, y_i) l'ensemble des coins détectés par R_{κ, σ, s_1} .
3. Soit $k = k + 1$, on cherche l'échelle maximisant le LoG (Laplacian of Gaussian) des points (x_i, y_i)

$$\hat{t} = \operatorname{argmin}_{t \in [0.7, \dots, 1.4]} |L_{xx}(x_i, y_i, ts_k) + L_{yy}(x_i, y_i, ts_k)|$$

4. $s_{k+1} = \hat{t}s_k$,
 (x_i, y_i) l'ensemble des coins détectés par $R_{\kappa, \sigma, s_{k+1}}$.
5. On itère 3-4 jusqu'à stabilisation du paramètre d'échelle s_k

Détection de blobs

- ▶ Lindeberg développe dans son livre une théorie structurale des images selon ces régions (qu'il appelle *blobs*).
- ▶ Donnée une échelle, essayons de décrire structurellement l'image :



(a) $s = 1$

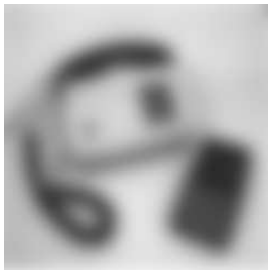
(b) seuillage



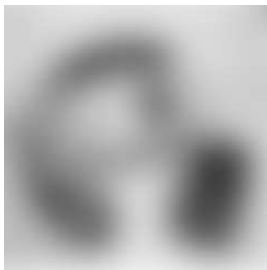
(c) $s = 2$



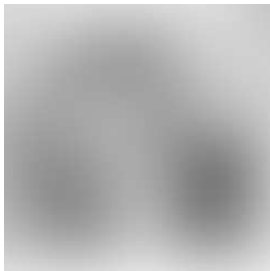
(d) $s = 4$



(e) $s = 8$



(f) $s = 16$



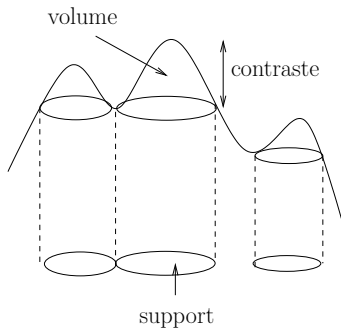
(g) $s = 32$



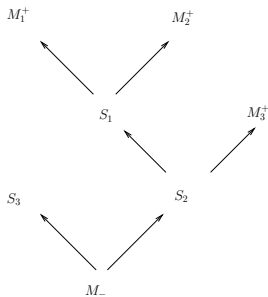
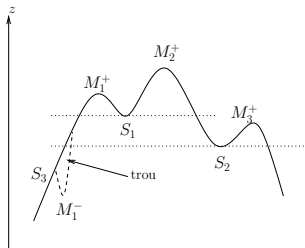
(h) $s = 64$

Définition d'un *blob* (Lindeberg)

- ▶ L'image est vue comme le graphe d'une surface
- ▶ Un *blob* est la portion de la surface comprise entre un maximum local et la courbe de niveau intersectant un point selle.



Structuration de l'image



- ▶ Pour déterminer les *blobs* et leur hiérarchie, on assimile le niveau de gris à une altitude que l'on fait croître.
- ▶ un changement de topologie des *blobs* indique un point selle,
- ▶ une apparition de *blob* indique un minimum local,
- ▶ une disparition de *blob* indique un maximum local.
- ▶ L'ordre d'apparition de ces points génère un arbre.

Structuration de l'image dans l'espace d'échelles

- ▶ Soit une représentation des *blobs* à l'échelle s , le calcul du laplacien fait à l'échelle s , permet de déterminer l'orientation des blobs :
 - ▶ $\nabla_s^2 I < 0$: convexe (*blob* blanc),
 - ▶ $\nabla_s^2 I > 0$: concave (*blob* noir).
- ▶ En faisant varier maintenant le facteur d'échelle : on obtient une description structurelle des blobs en 4D ($= (i, j, s, L(i, j, s))$).
- ▶ On peut alors étudier l'évolution dans l'espace des échelles de la hiérarchie des *blobs* et mettre en évidence les échelles où :
 - ▶ un *blob* apparaît,
 - ▶ deux *blobs* fusionnent,
 - ▶ un *blob* est découpé en deux,
 - ▶ un *blob* disparaît.

Détection de motifs invariant par changement d'échelle

Opérateur SIFT

- ▶ SIFT : Scale Invariant Feature Transform.
- ▶ Algorithme de David Lowe [Lowe, 1999, Lowe, 2004]
- ▶ Il s'agit de détecter des motifs dans les images.
- ▶ Représenter les éléments de l'image de façon à ne pas dépendre de l'échelle, de la position, de l'orientation, ... on pourra les comparer facilement à ce qu'on cherche.
- ▶ Il existe des représentations invariantes par translation (Transformée de Hough¹ par exemple), par rotation (détecteur de Harris), mais comment chercher une forme sans connaître à l'avance son échelle ?
- ▶ Principe de la méthode : calculer des différences de gaussiennes puis les orientations locales par histogrammes de gradients.

1. On peut rendre Hough invariant par échelle en payant le prix algorithmique

SIFT : étape 1

- ▶ Une façon d'être invariant par échelle est de calculer la représentation du motif à différentes échelles.
- ▶ Différence de gaussiennes :

$$D(I, \sigma_1, \sigma_2) = G_{\sigma_1} \star I - G_{\sigma_2} \star I, \quad \sigma_1 > \sigma_2$$

- ▶ Quelques propriétés :
 - ▶ Filtre passe-bande : élimination des structures de taille inférieure à σ_2 . Les deux gaussiennes préservent les structures de taille supérieure à σ_1 : la soustraction les éliminent.
 - ▶ Approximation du laplacien d'une convolution gaussienne (Laplacian of Gaussian) : puisque $L = G_{\sqrt{t}} \star I$ est solution de

$$\frac{\partial L}{\partial t} = \frac{1}{2} \nabla^2 L$$

on discrétise en temps le membre gauche et on obtient :

$$L(\mathbf{x}, t + dt) - L(\mathbf{x}, t) = D(I, t + dt, t) = \frac{1}{2} \nabla^2 L$$

Étape 1, suite

- ▶ Lowe propose de calculer les extrema de D : revient donc à faire de la détection de contours.
- ▶ Algorithme de Lowe :
 - ▶ Les échelles sont groupées par octave (l'octave est le double d'une échelle).
 - ▶ Par octave : un nombre fixé d'échelles.
 - ▶ On calcule $D(\mathbf{x}, \sigma) = L(\mathbf{x}, k\sigma) - L(\mathbf{x}, \sigma)$ avec k petit et fixé. On met en évidence les structures de taille σ .
 - ▶ En pratique, on calcule trois D :

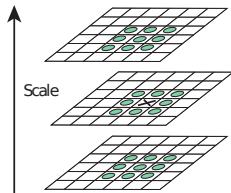
$$D(\mathbf{x}, k^{-1}\sigma) = L(\mathbf{x}, \sigma) - L(\mathbf{x}, k^{-1}\sigma)$$

$$D(\mathbf{x}, \sigma) = L(\mathbf{x}, k\sigma) - L(\mathbf{x}, \sigma)$$

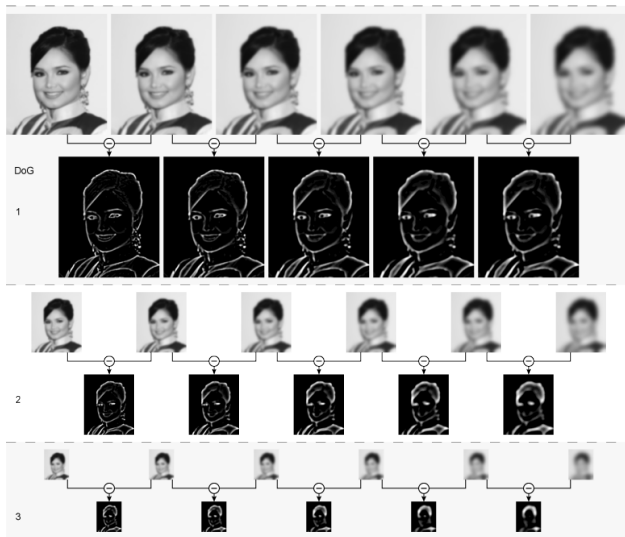
$$D(\mathbf{x}, k\sigma) = L(\mathbf{x}, k^2\sigma) - L(\mathbf{x}, k\sigma)$$

et on retient comme point d'intérêt les maxima locaux dans l'espace d'échelle (soit un 26 voisinage).

- ▶ octave suivante : on sous-échantillonne (pas 1,5) et on recommence.



Étape 1, suite



Étape 2 : relocalisation des extrema

- ▶ Le sous-échantillonnage pour les parties supérieures de la pyramide entraîne des erreurs de localisation des extrema des DoG.
- ▶ Une technique de relocalisation des extrema : D.L. ordre 2 dans l'espace d'échelles, au point d'intérêt \mathbf{x}_0 et dans un voisinage $\delta\mathbf{x} = (\delta x, \delta y, \delta\sigma)$:

$$D(\mathbf{x}_0 + \mathbf{x}) = D(\mathbf{x}_0) + \frac{\partial D}{\partial \mathbf{x}}(\mathbf{x}_0)^T \delta\mathbf{x} + \delta\mathbf{x}^T \frac{\partial^2 D}{\partial \mathbf{x}^2}(\mathbf{x}_0) \delta\mathbf{x} \quad (21)$$

- ▶ L'extrema est le point $\mathbf{x}_0 + \delta\mathbf{x}$ qui annule $\frac{dD(\mathbf{x}_0 + \delta\mathbf{x})}{d\delta\mathbf{x}}$. Donc :

$$\frac{\partial D}{\partial \mathbf{x}}(\mathbf{x}_0)^T + \frac{\partial^2 D}{\partial \mathbf{x}^2}(\mathbf{x}_0) \delta\mathbf{x} = 0$$

- ▶ On localise l'extrema avec :

$$\delta\mathbf{x} = \left(\frac{\partial^2 D}{\partial \mathbf{x}^2}(\mathbf{x}_0) \right)^{-1} \frac{\partial D}{\partial \mathbf{x}}(\mathbf{x}_0) \quad (22)$$

- ▶ Si $|\delta\mathbf{x}| > 0.5$, on prend comme point $\mathbf{x}_0 + \delta\mathbf{x}$.

Étape 3 : élimination des points non pertinents

- ▶ En pratique, on trouve beaucoup (trop) de maxima locaux dans l'espace d'échelle et beaucoup ne sont pas pertinents.
- ▶ Remplaçons l'éq. (22) dans (21) :

$$D(\mathbf{x}_0 + \delta\mathbf{x}) = D(\mathbf{x}_0) + \frac{1}{2} \frac{\partial D}{\partial \mathbf{x}}(\mathbf{x}_0)^T \delta\mathbf{x}$$

- ▶ Si $|D(\mathbf{x}_0 + \delta\mathbf{x})| < 3\%$: on élimine ce point (de contraste trop faible)

Étape 3 : élimination des points non pertinents

- ▶ Les points de contours : ils répondent fortement mais sont par nature instable (peuvent se déplacer le long des arêtes selon les échelles). Ils sont identifiés par un critère de type Harris.
- ▶ On calcule le hessien de D :

$$H = \begin{pmatrix} \frac{\partial^2 D}{\partial x^2} & \frac{\partial^2 D}{\partial x \partial y} \\ \frac{\partial^2 D}{\partial x \partial y} & \frac{\partial^2 D}{\partial y^2} \end{pmatrix}$$

- ▶ Les deux valeurs propres de $H(\mathbf{x}_0)$ donnent les valeurs des deux courbures principales au point \mathbf{x}_0 de la surface $\mathbf{x} \mapsto D(\mathbf{x})$.
- ▶ Un point de contours possède une courbe principale significativement plus grande que l'autre. Critère :

$$R = \frac{\text{trace}(H)^2}{\det(H)} = \frac{(r+1)^2}{r}, \quad r = \frac{\tau_1}{\tau_2}$$

- ▶ On garde les points quels que $R(\mathbf{x}_0) < 10$.

Étape 4 : orientation des points d'intérêt

- ▶ On calcule leur orientation pour garantir par la suite une représentation invariante par rotation.
- ▶ Orientation : en chaque point du voisinage du point d'intérêt $(\mathbf{x}_0 = (x \ y \ \sigma)^T$, on calcule l'orientation et la magnitude du vecteur gradient de l'image lissée $L = I \star G_\sigma$ (à l'échelle du point d'intérêt) :

$$M(x, y) = \sqrt{(L_{x+1,y} - L_{x-1,y})^2 + (L_{x,y+1} - L_{x,y-1})^2}$$
$$\theta(x, y) = \tan^{-1} ((L_{x,y+1} - L_{x,y-1}) / ((L_{x+1,y} - L_{x-1,y})))$$

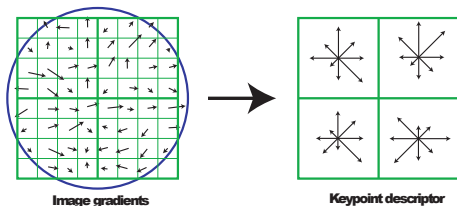
- ▶ On forme l'histogramme de ces orientations (regroupées par pas de 10°).
- ▶ Chaque entrée dans l'histogramme est pondérée par la magnitude et un poids gaussien d'écart-type 1.5σ .
- ▶ Les fréquences les plus élevées correspondent aux orientations dominantes : on garde celles qui représentent 80% des valeurs.

Le descripteur SIFT

- ▶ Un point d'intérêt est donc le quadruplet (x, y, σ, θ) tel que :
 - ▶ (x, y, σ) est un max. local dans l'espace d'échelle et $|D(x, y, \sigma)| > 3\%$ et $|R(x, y, \sigma)| < 10$.
 - ▶ $\theta(x, y, \sigma)$ est une orientation dominante.
- ▶ Il peut y avoir plusieurs orientations au même point de l'espace d'échelle
- ▶ Il peut aussi y avoir plusieurs échelles au même point du domaine image.

Le descripteur SIFT

- ▶ Le descripteur SIFT : a un point d'intérêt donnée (position-échelle-orientation) on associe une série de valeurs (le descripteur) qui caractérise ce point :
 - ▶ On considère un voisinage 16×16 : on calcule sur chaque quadrant un histogramme d'orientation (même algo que précédemment) : les orientations sont calculées par rapport à l'orientation du point : invariance par orientation.



- ▶ Un motif sera caractérisé par ces histogrammes d'orientations autour de ses points d'intérêts.
- ▶ Couplé à des techniques de reconnaissance de formes, on peut alors rechercher ces motifs (descripteurs) dans une image.

Application

Reconnaissance d'objets



Application

Reconnaissance d'objets



Application

Reconnaissance d'objets



Sélection automatique de l'échelle

- ▶ On l'a vu : les espaces d'échelles linéaires détectent les objets dont la taille est voisine de l'échelle : il faut connaître à l'avance cette échelle !
- ▶ Représentons une structure 1D par le signal suivant :

$$f(x) = \sin \omega_0 x$$

ω_0 , la pulsation, représente une structure de période $\frac{1}{\omega_0}$.

- ▶ Représentation en espace d'échelle de ce signal :

$$\begin{aligned} L(x; t) &= f \star g_{\sqrt{t}}(x) \\ &= e^{-\omega_0^2 t/2} \sin \omega_0 x \end{aligned}$$

- ▶ On a : $|L(\cdot, t)|_\infty = e^{-\omega_0^2 t/2}$: monotone décroissant en t .

Sélection automatique de l'échelle

- Pour le signal dérivé, on a :

$$L^{(m)}(x, t) = \frac{\partial^m L(x, t)}{\partial x^m}$$
$$|L^{(m)}(., t)|_{\infty} = \omega_0^m e^{-\omega_0^2 t/2}$$

Le maximum est monotone décroissant dans l'espace d'échelle.

- Il n'y a donc pas d'échelle idéale pour la détection de la structure de taille $\frac{1}{\omega_0}$.
- [Lindeberg, 1998] : introduisons la *dérivée normalisée* :

$$\partial_x^\gamma = t^{\gamma/2} \frac{\partial}{\partial x}, \quad \gamma > 0$$

Sélection automatique de l'échelle

- ▶ Calculons la m -ième dérivée normalisée de L :

$$L_{\gamma}^{(m)}(x, t) = \left(t^{\gamma/2}\right)^m \frac{\partial^m L}{\partial x^m}$$

- ▶ Le maximum du signal dérivée est alors :

$$|L_{\gamma}^{(m)}(\cdot, t)|_{\infty} = t^{m\gamma/2} \omega_0^m e^{-\omega_0^2 t/2}$$

- ▶ Le maximum n'est plus monotone dans l'espace des échelles : il possède un unique maximum, c-à-d qu'il existe une échelle pour laquelle la réponse est maximale. Cette échelle est fonction de ω_0 :

$$t_{\max} = \frac{\gamma m}{\omega_0^2}$$

- ▶ Ce maximum est fonction de ω_0 :

$$\max(|L_{\gamma}^{(m)}(\cdot, t)|_{\infty}) = \frac{(\gamma m)^{\gamma m/2}}{e^{\gamma m/2}} \omega_0^{(1-\gamma)m} \quad (23)$$

Sélection automatique de l'échelle

- En pratique : on peut prendre $\gamma = 1$, cas particulier pour lequel le maximum du signal dérivé est indépendant de la taille de la structure (voir éq. (23)).

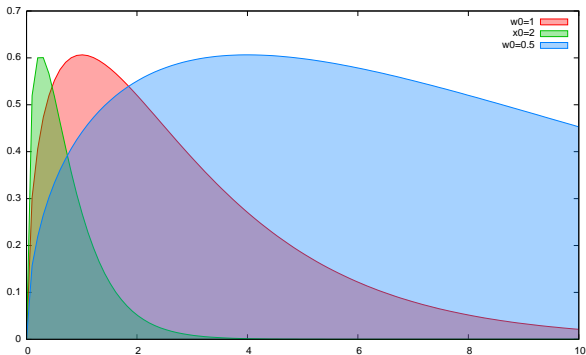


Figure : Amplitude maximale de $L_\gamma^{(m)}$

Sélection automatique de l'échelle

► Algorithme :

1. Calculer $\partial_{xx}^\gamma L$, $\partial_{yy}^\gamma L$, $\partial_{xy}^\gamma L$ pour différentes échelles t ($t = 3, 4, 5, \dots$).
En pratique, on prendra $\gamma = 1$, et on a donc $\partial_{xx}^\gamma L = tL_{xx}$ (idem pour $\partial_{yy}^\gamma L$ et $\partial_{xy}^\gamma L$)
2. Détection de *blob* : calculer

$$F(t) = \text{trace } H(L) = t(L_{xx} + L_{yy})$$

ou bien

$$F(t) = \det H(L) = t(L_{xx}L_{yy} - L_{xy}^2)$$

pour les différentes échelles.

3. Chercher les maxima locaux de F **à la fois en espace et en échelle** : la valeur de l'échelle renseigne sur la taille du *blob*.

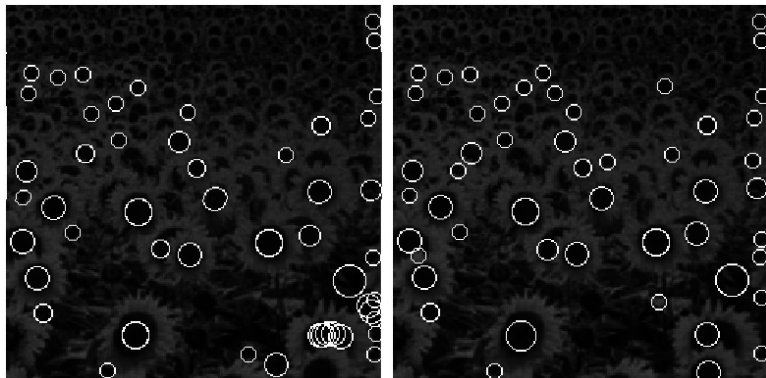
Sélection automatique de l'échelle



Figure : Champ de tournesols

Applications - Sélection automatique de l'échelle


Sélection automatique de l'échelle




(a) Trace du Hessien


(b) Déterminant du Hessien


Figure : Les diamètres des cercles sont calculés à partir de l'échelle trouvée (merci à Hailin Shi).







 Burt, P. J. (1981).
Fast filters transforms for image processing.
Computer Vision and Image Processing, 16 :20–51.

 Canny, J. (1986).
A computational approach to edge detection.
IEEE Trans. on PAMI, 8(6) :679–698.

 Crowley, J. L. (1981).
A Representation for Visual Information.
PhD thesis, Carnegie-Mellon University, Robotics Institute,
Pittsburgh, Pennsylvania.

 Klinger, A. (1971).
Pattern and search statistics.
Optimizing Methods in Statistics.

 Koenderink, J. J. (1984).
The structure of images.
Biological Cybernetics, 50 :363–370.

-  Lindeberg, T. (1994).
Scale-Space Theory in Computer Vision.
Kluwer Academic Publishers.
-  Lindeberg, T. (1998).
Feature detection with automatic scale selection.
International Journal on Computer Vision, 30(2) :77–116.
-  Lowe, D. (1999).
Object recognition from local scale-invariant features.
In International Conference on Computer Vision, pages 1–8.
-  Lowe, D. (2004).
Distinctive image features from scale-invariant keypoints.
International Journal on Computer Vision, 60 :91–110.
-  Marr, D. and Hildreth, E. (1980).
Theory of edge detection.
Proc. of the Royal Society of London, B207 :187–217.
-  Rosenfeld, A. and Thurston, M. (1971).

Edge and curve detection for visual scene analysis.

IEEE Transaction on Computers, 20(5) :563–569.



Uhr, L. (1972).

Layered 'recognition cone' networks that preprocess, classify and describe.

IEEE Trans. Comput., pages 759–768.



Witkin, A. P. (1983).

Scale-space filtering.

In *Proc. 8th Int. Joint Conf. Art. Intell.*, pages 1019–1022,
Karlsruhe, Germany.

Article fondateur sur les espaces d'échelles (1D).