

Chapitre 3

Caractéristiques multi-échelles

3.1 Introduction

Dans ce chapitre on s'intéresse à la caractérisation locale de la géométrie dans les images, un élément primordial de tous les algorithmes de vision par ordinateur. De façon générale, cette caractérisation consiste à évaluer tous les pixels de l'image par rapport à une mesure géométrique, puis éventuellement à sélectionner un sous-ensemble des pixels les plus significatifs vis-à-vis de cette mesure.

Pour caractériser localement la géométrie dans une image, il est nécessaire de préciser ce qu'on entend par «localement», c'est-à-dire d'indiquer une *échelle* d'analyse. En effet, les objets réels, la réalité physique, par opposition aux objets mathématiques, existent sous forme d'entités bien différentes selon l'échelle à laquelle on les considère : Une feuille, un arbre, une forêt présentent ainsi des caractéristiques de forme, de couleur, de texture radicalement différentes. Etudier le comportement local d'un objet à partir de dérivées estimées par différences finies comme dans le chapitre précédent ne permettra d'appréhender qu'une petite partie de la réalité physique représentée par l'image : l'échelle la plus fine (celle du pixel).

La théorie des espaces d'échelle [50] est un modèle qui est apparu progressivement dans le domaine de la vision artificielle, principalement pour le bas niveau (i.e. traitement d'images), pour prendre en compte cette nature foncièrement multiéchelles des données 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 en général pas connues à l'avance. C'est pourquoi il est essentiel que les opérateurs considèrent l'image à plusieurs échelles ; c'est à ce titre que les espaces d'échelles constituent un modèle formel adapté.

La section 3.2 introduit de façon générale les espaces d'échelles linéaires. Les sections suivantes montrent l'utilisation de l'analyse multi-échelles pour le calcul de caractéristiques très importantes en pratique : les grandeurs différentielles, les contours et les points d'intérêt. La section 3.6 présente quelques exemples d'espaces d'échelles non linéaires.

3.2 Espace d'échelles linéaire

Les opérateurs de traitement d'images dont le support de calcul est réduit au pixel et à ses voisins immédiats - tels que les filtres 3×3 présentés au chapitre précédent pour approximer les grandeurs différentielles - permettent d'estimer des structures à l'échelle la plus fine : celle du pixel. Si l'on souhaite estimer, avec le même opérateur, une structure à différentes échelles, il faut opérer sur l'image à des résolutions différentes. C'est le principe de la *multi-résolution*, qui consiste à produire une suite d'images de résolution décroissante, ou *pyramide*, et à appliquer des traitements locaux à différents niveaux de résolution [21].

Ainsi, sur la figure 3.1, la première ligne montre le calcul d'une dérivée première verticale par un opérateur local calculé à la résolution 256×256 (niveau 0 de la pyramide). La seconde ligne montre le même calcul réalisé sur la même image à la résolution 16×16 , obtenue en divisant par 16 le nombre d'échantillons spatiaux dans chaque dimension (niveau 4 de la pyramide dyadique, pour laquelle chaque dimension de l'image est divisée par 2 en passant d'un niveau

au supérieur).

Mais, comme on l'a vu dans le premier chapitre, la pyramide ne doit être obtenue par un simple sous-échantillonnage, car alors les artefacts dus à l'aliasing seraient naturellement détectés par les opérateurs locaux dans les résolutions plus basses. La sous-résolution doit donc combiner un filtrage spatial en préalable au sous-échantillonnage. On a vu (Chapitre 1), que dans ces bonnes conditions, l'image sous-échantillonnée était finalement équivalente à l'image filtrée à pleine résolution.

C'est précisément sur cette équivalence que repose la représentation *multi-échelles*. Au contraire de la précédente, elle n'implique pas de sous-résolution des images, mais elle identifie l'indice d'échelle, significatif de la taille des structures concernées par l'opérateur, avec le paramètre de portée spatial d'un filtre caractérisant cet opérateur. Ainsi, sur la figure 3.1, la dernière ligne montre, à gauche l'image à l'échelle 16, au sens du filtre passe-bas idéal ($u_M = v_M = 8$). Du point de vue information, cette image est équivalente à l'image 16×16 au dessus. À droite, on peut voir l'estimation multi-échelle du gradient vertical, correspondant ici à une différence locale calculée sur l'image de gauche.

Dans l'axiomatique des espaces d'échelles, une image I est donc associée à un paramètre d'échelle s , lui-même identifié à l'indice de portée spatial d'un opérateur P :

$$I_s = P_s(I) \quad (3.1)$$

Pour les espaces d'échelles linéaires, P sera supposé linéaire et invariant en translation, c'est-à-dire une convolution :

$$I_s = I \star h_s \quad (3.2)$$

L'axiomatique des espaces d'échelles linéaires est fondée sur un certain nombre de propriétés exprimant l'absence d'*a priori* quelconque sur les images :

- invariance en translation
- invariance en rotation
- invariance en échelle

Pour des raisons calculatoires, on impose aussi que la composition de 2 opérateurs de changement d'échelle produise un changement d'échelle réalisable avec un seul opérateur. Cela correspond à la propriété dite de semi-groupe, alléguant l'existence d'une loi \boxplus opérant sur les indices d'échelles, telle que :

$$h_{s_1} \star h_{s_2} = h_{s_1 \boxplus s_2} \quad (3.3)$$

Lindeberg et ter Haar Romeny [44] (Chapitre 1) ont montré que les propriétés d'invariance et de semi-groupe impliquent que l'opérateur de changement d'échelle h_s coïncide avec le noyau gaussien de variance $2s$:

$$h_s = G_{\sqrt{2s}} = \frac{1}{4\pi s} e^{-\frac{(x^2+y^2)}{4s}} \quad (3.4)$$

Les espaces d'échelles linéaires sont donc généralement synonymes d'espaces d'échelles gaussiens. La figure 3.2 montre ainsi la représentation d'un signal 1d dans l'espace d'échelle gaussien, sous la forme d'un faisceau de courbes, correspondant à l'application de filtres de convolution gaussien de variance croissante.

Cette figure illustre l'idée clef des espaces d'échelles, qui est le *principe de causalité* [23], selon lequel le signal ne doit pas faire apparaître de nouvelles

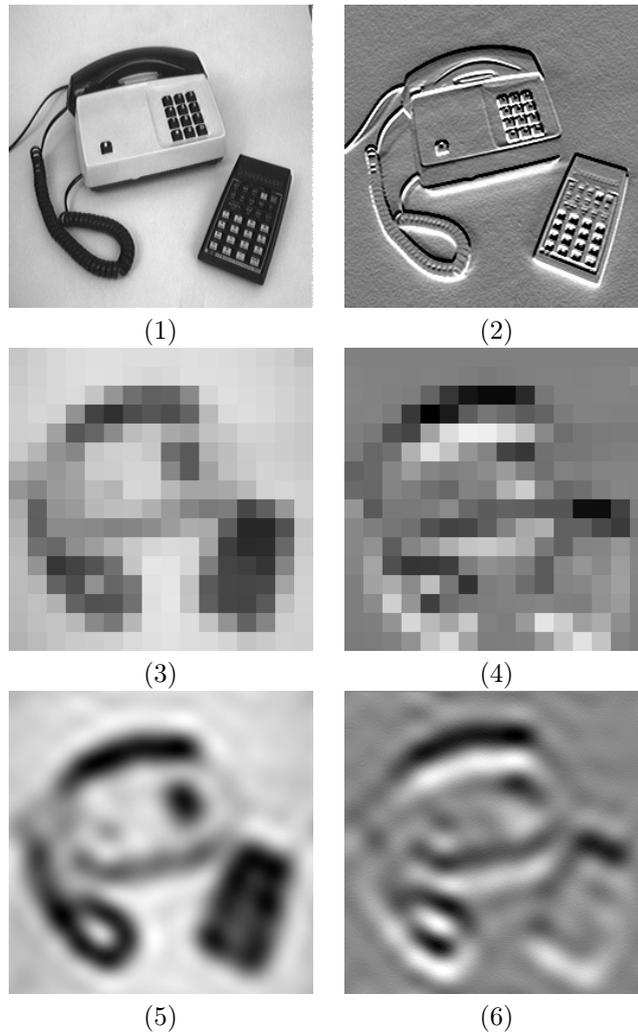


FIGURE 3.1 – Multi-résolution ou Multi-échelle (1) et (2) : une image (256×256), et l'estimée de sa dérivée verticale à l'échelle du pixel (Sobel vertical). (3) et (4) : estimation de la dérivée verticale en multi-résolution, à l'échelle 16 (4ème niveau de la pyramide dyadique). (5) et (6) : estimation de la dérivée verticale en multi-échelle, à l'échelle 16.

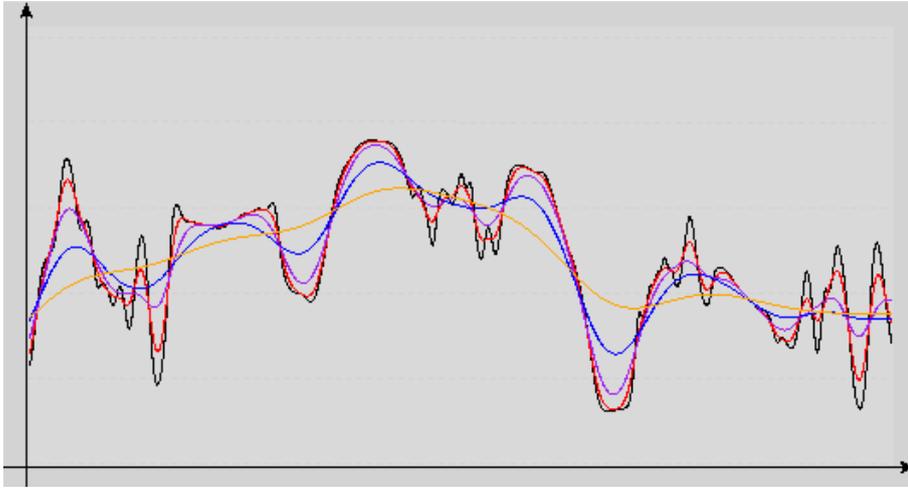


FIGURE 3.2 – Représentation multi-échelles d’un signal 1d discret de 512 échantillons, pour les écarts-type $\sigma \in \{2, 4, 8, 16, 32\}$, correspondant aux échelles $s \in \{2, 8, 32, 128, 512\}$.

structures lorsque l’échelle augmente. Pour un faisceau de courbes 1d, ce principe se traduit par :

- pas d’apparition de nouveaux extrema locaux.
- pas d’amplification des extrema existants (i.e. lorsque l’échelle augmente, la valeur d’un maximum diminue et la valeur d’un minimum augmente).

En 2d, cette propriété n’est plus vraie : Si l’on voit l’image comme un relief, où la hauteur correspond au niveau de gris, il suffit d’imaginer 2 cols de hauteurs différentes reliés par une corniche plus fine que le support spatial des deux cols. A l’échelle la plus fine, cette «image» ne comporte qu’un maximum local, qui est le sommet du plus grand col. Cependant, le lissage gaussien fait progressivement disparaître la corniche, si bien qu’à une certaine échelle, le deuxième col apparaît comme un nouveau maximum local.

Malgré cette limitation inhérente au filtrage linéaire en 2d, le noyau gaussien reste le meilleur support de calcul de l’espace d’échelle linéaire. Une de ses applications les plus importantes est l’estimation des dérivées partielles, qu’on traite dans la section suivante.

3.3 Dérivées multi-échelles

Un des intérêts fondamentaux de l’espace d’échelle linéaire réside dans le fait que la convolution commute avec la dérivation :

$$\partial^n (I \star h_s) = I \star (\partial^n h_s) \quad (3.5)$$

Ainsi, dans le cadre gaussien, estimer une dérivée à l’échelle s revient à calculer une convolution avec la dérivée de la gaussienne d’écart type $\sigma = \sqrt{2s}$. Soit G_σ la fonction gaussienne 2d d’écart-type σ , on note $G_\sigma^{(i,j)}$ sa dérivée d’ordre (i, j) :

$$G_\sigma^{(i,j)} = \frac{\partial^{i+j}}{\partial x^i \partial y^j} G_\sigma \quad (3.6)$$

Le fondement du calcul différentiel dans les images (comme dans tout signal multidimensionnel discrétisée de façon régulière), repose donc sur la définition suivante :

$$\left(\frac{\partial^{i+j} I}{\partial x^i \partial y^j} \right)_s \stackrel{\text{définition}}{=} I \star G_{\frac{\sigma}{\sqrt{2s}}}^{(i,j)} \quad (3.7)$$

Cela signifie qu'une dérivée n'a de sens qu'estimée à une certaine échelle, en imposant à l'image discrète un niveau de régularité en relation avec la portée spatiale de la convolution. La figure 3.3 montre le calcul des dérivées jusqu'à l'ordre 2 pour 3 échelles différentes, et les noyaux de convolution correspondants. Dans la suite on notera couramment $I_{x^i y^j} = \frac{\partial^{i+j} I}{\partial x^i \partial y^j}$, en omettant l'échelle d'estimation.

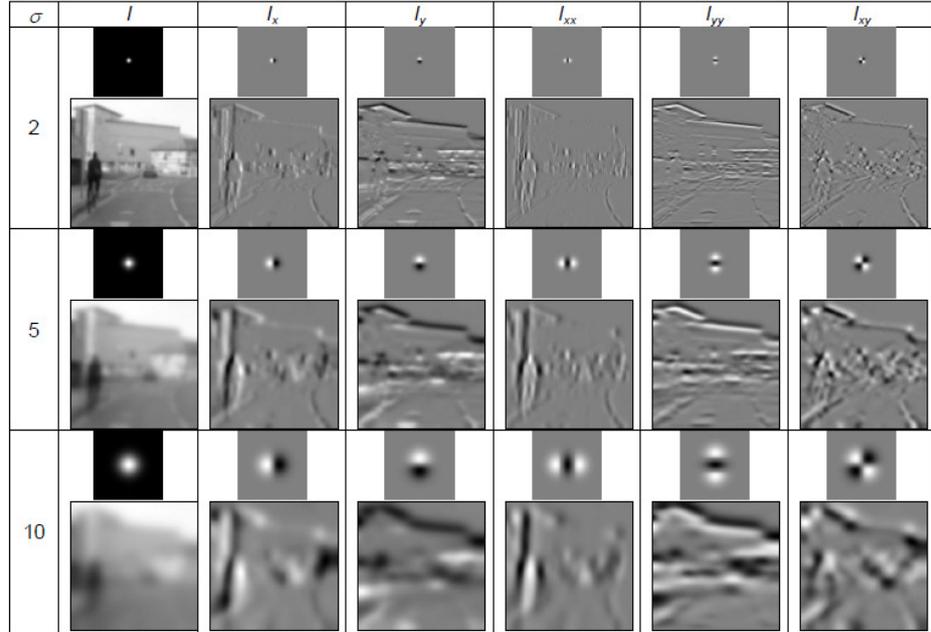


FIGURE 3.3 – Les dérivées à l'ordre 0, 1 et 2, pour 3 échelles différentes $s \in \{2, 12.5, 50\}$, soit $\sigma \in \{2, 5, 10\}$. Au dessus de chaque image (200×200), on peut voir la réponse impulsionnelle de la (dérivée de) gaussienne $G_\sigma^{(i,j)}$ correspondante (image 100 \times 100).

A partir de ces dérivées, on peut calculer un grand nombre de caractéristiques locales dans les images, à commencer par les mesures classiques de la géométrie différentielle. Ainsi à l'ordre 1, on s'intéresse au gradient $\nabla I = (I_x, I_y)$, qui induit un repère local (g, t) à chaque pixel, où $g = \frac{\nabla I}{\|\nabla I\|}$ est la direction du gradient et $t = g^\perp$ la direction de l'isophote. On a $I_g = \|\nabla I\|$ et $I_t = 0$. La figure 3.4 illustre ces grandeurs sur une image, pour 3 échelles différentes.

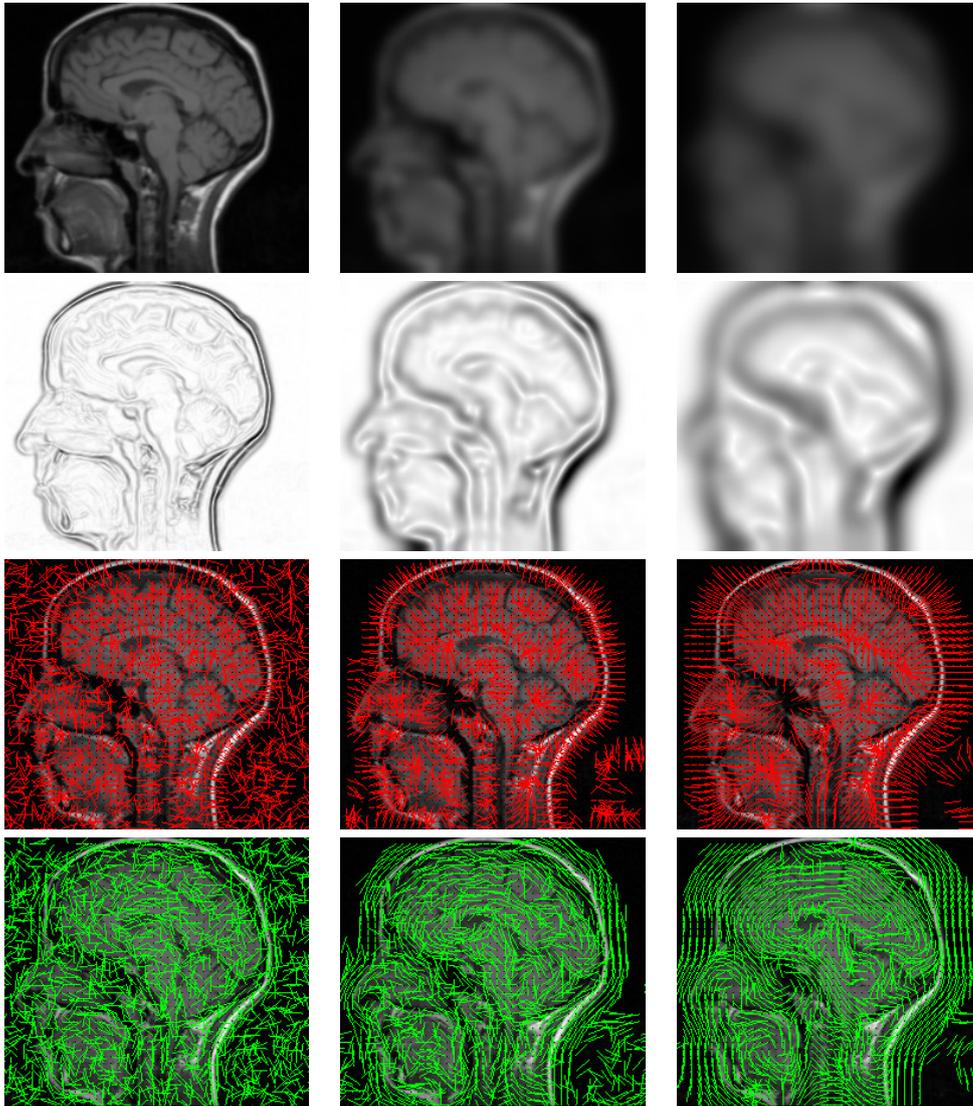


FIGURE 3.4 – Grandeurs différentielles d'ordre 1. De gauche à droite : échelles $s = 0.5$, $s = 12.5$, et $s = 50.0$. De haut en bas : I , $I_g = \|\nabla I\|$, Direction du gradient g , Direction de l'isophote t

A l'ordre 2, la matrice hessienne $H_I = \begin{pmatrix} I_{xx} & I_{xy} \\ I_{yy} & I_{xy} \end{pmatrix}$ définit les courbures principales dans l'image, par ses vecteurs et valeurs propres. On note λ_I et Λ_I les 2 valeurs propres de H_I , telles que $\lambda_I \leq \Lambda_I$. On a :

$$\lambda_I = \beta_I - \alpha_I \quad (3.8)$$

$$\Lambda_I = \beta_I + \alpha_I \quad (3.9)$$

avec

$$\beta_I = \frac{1}{2}(I_{xx} + I_{yy}) \quad (3.10)$$

$$\alpha_I = \frac{1}{2}\sqrt{(I_{xx} - I_{yy})^2 + 4I_{xy}^2} \quad (3.11)$$

Les dérivées secondes mesurent la courbure de l'image, c'est-à-dire le niveau de «pincement» du relief. La direction des courbures principales est donnée par les vecteurs propres de la matrice hessienne, qui sont orthogonaux (voir Fig. 3.6). L'intensité de la courbure, qui peut être négative ou positive, est donnée par les valeurs propres (voir Fig. 3.5).

D'autres mesures d'ordre 2 jouent un rôle important dans la caractérisation locales des images :

$$\begin{aligned} \Delta_I &= \lambda_I + \Lambda_I = \frac{I_{xx} + I_{yy}}{2} \\ \|H_I\|_F &= \sqrt{\lambda_I^2 + \Lambda_I^2} = \sqrt{I_{xx}^2 + I_{yy}^2 + 2I_{xy}^2} \\ \det(H_I) &= \lambda_I \Lambda_I = I_{xx}I_{yy} - I_{xy}^2 \end{aligned}$$

- Le *laplacien* Δ_I , égal à la trace de la matrice hessienne, représente une sorte de courbure moyenne (Fig. 3.7).
- La *norme de Frobenius* de la matrice hessienne $\|H_I\|_F$, représente l'intensité totale de la courbure (Fig. 3.6).
- Le *déterminant* de la matrice hessienne $\det(H_I)$ identifie les spots circulaires dans les images, pour lesquels la valeur absolue de la courbure est forte dans toutes les directions (Fig. 3.7).

On peut remarquer sur les images de la figure 3.5 que le signe des valeurs propres détermine la nature de la courbure. Lorsque les valeurs propres sont de même signe (par exemple les spots sombres sur la 2e ligne ou les spots clairs sur la 3e ligne de la figure 3.5), on est en présence d'une structure de type elliptique ou blob. Lorsqu'elles sont de signes différents (par exemple les lignes claires sur la 2e ligne ou les lignes sombres sur la 3e ligne de la figure 3.5), on est en présence d'une structure de type tubulaire ou crête.

Les dérivées jusqu'à l'ordre 2 fournissent ainsi une catégorisation de la géométrie locale selon l'ordre de dérivation dominant et le signe des valeurs propres de la matrice hessienne. La table 3.1 présente cette catégorisation en 4 classes (6 si on tient compte de la polarité de la courbure). Les imageries qui apparaissent dans le tableau ont été produites à partir de la formule de Taylor calculée dans un voisinage 15×15 à partir des dérivées partielles obtenues dans une image naturelle, pour 6 pixels appartenant à chacune des catégories.

Enfin, d'autres dérivées secondes jouent un rôle essentielle en traitement d'images, ce sont les courbures dans les direction du repère local au premier ordre

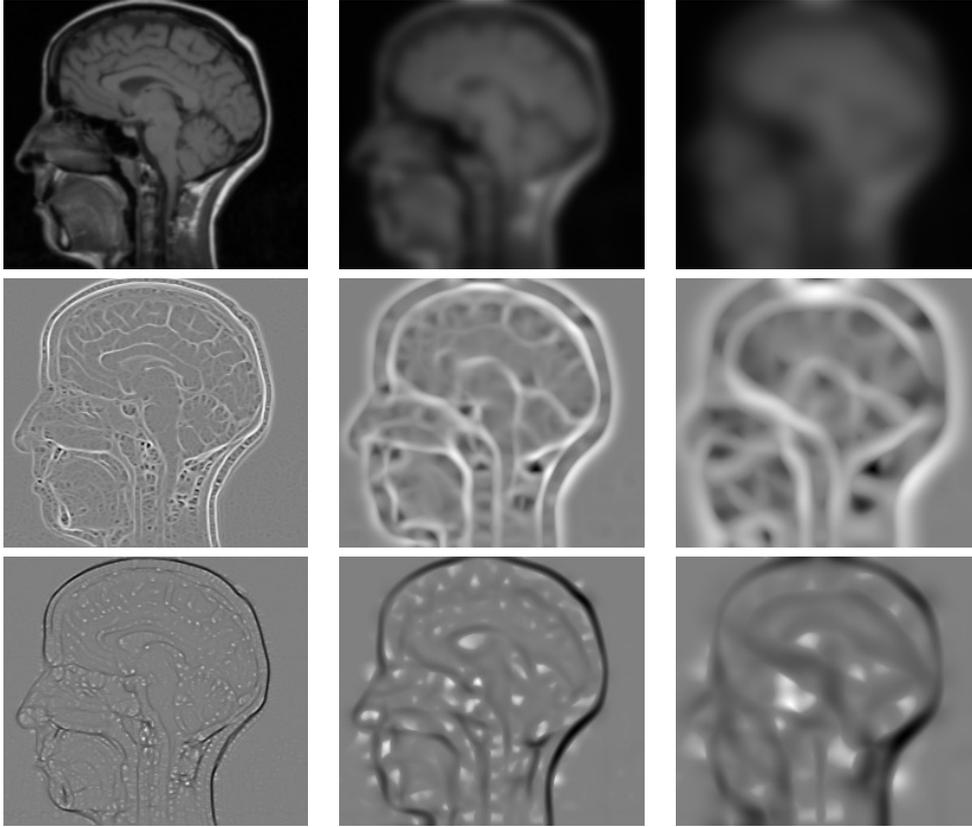
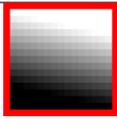
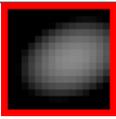
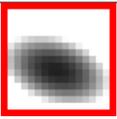
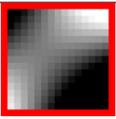
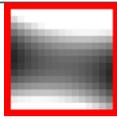


FIGURE 3.5 – Grandeurs différentielles d'ordre 2 : les courbures principales. De gauche à droite : échelles $s = 0.5$, $s = 12.5$, et $s = 50.0$. De haut en bas : I , la plus grande valeur propre de la hessienne Λ_I , et la plus petite valeur propre λ_I .

TABLE 3.1 – Catégorisation locale des pixels en fonction de l'ordre dominant de dérivation. À l'ordre 2, la valeur propre qui apparaît en gras est celle de plus grande valeur absolue, correspondant à la courbure dominante.

0 : Plateau	1 : Rampe	2 : Courbure			
$\ \nabla_I\ \simeq 0$ $\ H_I\ _F \simeq 0$	$\ \nabla_I\ \gg 0$ $\ H_I\ _F \simeq 0$	elliptique : $\Lambda_I \lambda_I > 0$		$\ H_I\ _F \gg 0$ selle (resp. tubulaire) : $\Lambda_I \lambda_I < 0$	
					
		$\Lambda_I < 0$ $\lambda_I < 0$	$\Lambda_I > 0$ $\lambda_I > 0$	$\Lambda_I > 0$ (resp. $\simeq 0$) $\lambda_I < 0$	$\Lambda_I > 0$ $\lambda_I < 0$ (resp. $\simeq 0$)

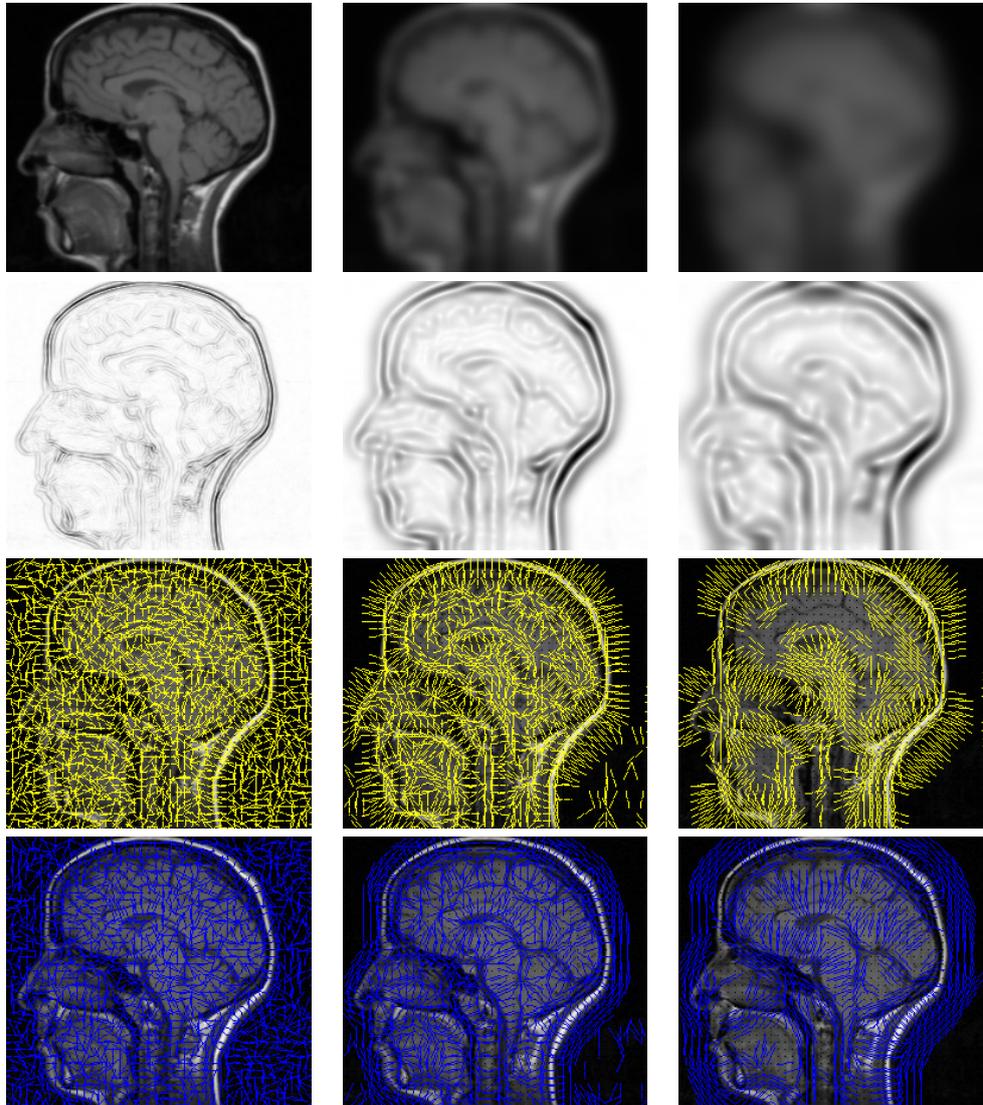


FIGURE 3.6 – Grandeurs différentielles d'ordre 2. De gauche à droite : échelles $s = 0.5$, $s = 12.5$, et $s = 50.0$. De haut en bas : I , $\|H_I\|_F$, la norme de Frobenius de la matrice hessienne, la direction du premier vecteur propre (associé à la grande valeur propre Λ_I), et la direction du second vecteur propre (associé à la petite valeur propre λ_I)

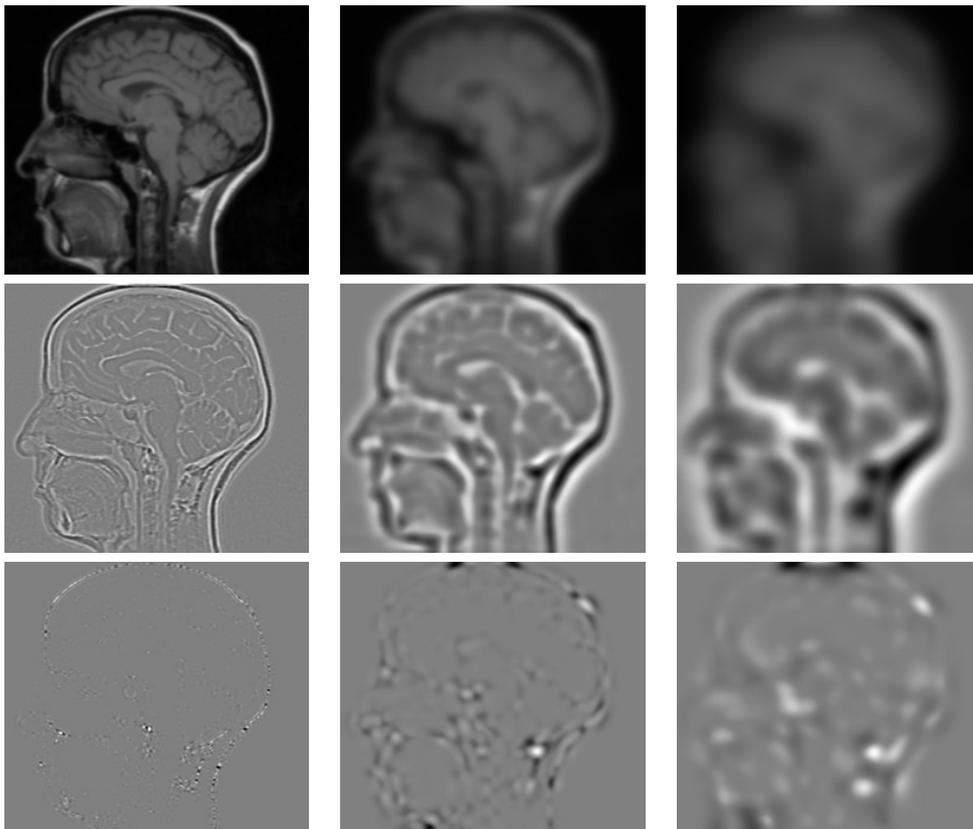


FIGURE 3.7 – Grandeurs différentielles d'ordre 2 : Autres mesures hessiennes. De gauche à droite : échelles $s = 0.5$, $s = 12.5$, et $s = 50.0$. De haut en bas : I , ΔI , le laplacien, et $\det(H_I)$, le déterminant de la matrice hessienne.

(g, t) défini par le gradient et l'isophote. La dérivée seconde dans la direction du gradient est utilisée pour le calcul des contours (voir section suivante), et vaut :

$$I_{gg} = g^\top \cdot H_I \cdot g \quad (3.12)$$

$$= \frac{I_{xx}I_x^2 + 2I_{xy}I_xI_y + I_{yy}I_y^2}{I_x^2 + I_y^2} \quad (3.13)$$

De même la dérivée seconde dans la direction de l'isophote vaut :

$$I_{tt} = \frac{I_{xx}I_y^2 - 2I_{xy}I_xI_y + I_{yy}I_x^2}{I_x^2 + I_y^2} \quad (3.14)$$

Cette grandeur est liée à la courbure de la ligne de niveau (du contour) qui vaut $\kappa_I = -\frac{I_{tt}}{\|\nabla I\|}$, avec $R_I = \frac{1}{|\kappa_I|}$, où R_I est le rayon du cercle osculateur au contour. Cette égalité sera démontrée dans un prochain chapitre.

3.4 Contours

La détection de contours est l'un des premiers problèmes «historique» du traitement d'images. Il s'agit de transformer l'image en un ensemble de courbes définissant les «frontières significatives». Ces courbes représentent une simplification drastique du contenu de l'image, et on peut en extraire des descripteurs utiles dans plusieurs applications.

Il serait abusif de dire que ce problème est réglé aujourd'hui, car aucun détecteur ne fournit une réponse parfaitement adaptée à une application donnée, et encore moins conforme à une notion de frontière sémantique entre objets. Néanmoins la théorie des espaces d'échelles permet de définir très proprement les détecteurs de contours, tout en fournissant des algorithmes pratiques très simples à utiliser [6], [44] (Chapitre 2).

Sur un signal 1d dérivable deux fois, la notion de contours peut être définie comme le lieu des maxima locaux de la dérivée première, ou, ce qui est équivalent, le lieu des passage par zéro de la dérivée seconde (voir Figure 3.8).

Pour retrouver cette notion naturelle en 2d, il faut se placer, en chaque point, dans la direction de variation principale de l'image, c'est-à-dire la direction du gradient. Si l'on considère l'image I au voisinage d'un pixel de contour (i, j) , en observant le signal dans la direction et le sens du gradient $\nabla I(i, j)$ (voir Figure 3.9), alors on retrouve une fonction f 1d, de type échelon, comme illustré Figure 3.8. Les contours dans les images sont donc définis comme les lieux de maxima locaux du module du gradient, dans la direction du gradient, ou encore, les lieux de passage par zéro de la dérivée seconde de l'image dans la direction du gradient. Etant, de plus, dans le sens du gradient positif, on aura toujours un échelon croissant, et donc une dérivée troisième négative (voir Figure 3.8).

Rappelons nos notations des dérivées partielles de l'image I : $I_x = \frac{\partial I}{\partial x}$, $I_y = \frac{\partial I}{\partial y}$, $I_{xx} = \frac{\partial^2 I}{\partial x^2}$, $I_{xy} = \frac{\partial^2 I}{\partial x \partial y}$, et ainsi de suite. On va donc se placer dans le repère local (g, t) défini par le gradient (voir Figure 3.9).

Pour calculer les dérivées d'ordre 1 à 3 selon g et t , on applique la règle de dérivation des fonctions composées, pour toute fonction v :

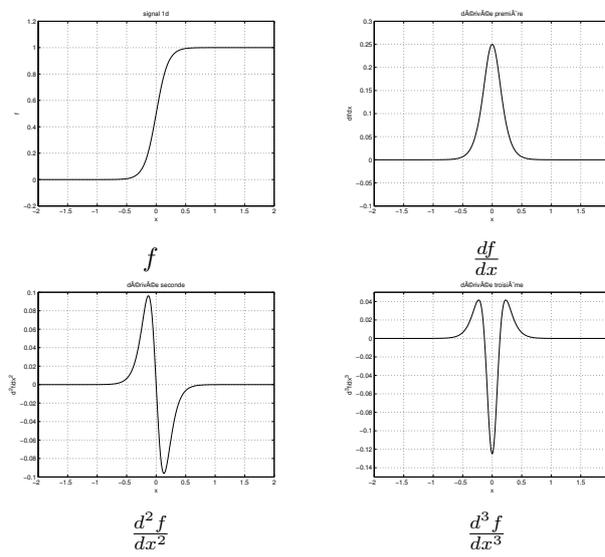


FIGURE 3.8 – Un contour 1d (type échelon) en $x = 0$, et ses dérivées d'ordre 1 à 3.

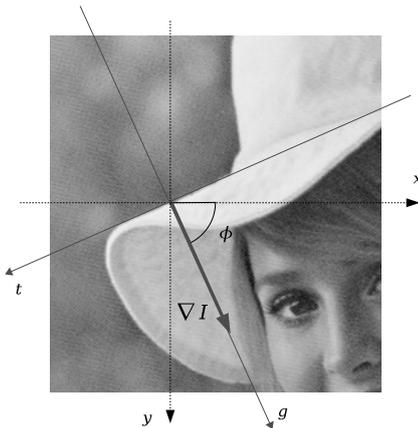


FIGURE 3.9 – Le repère (g, t) du gradient local.

$$\frac{\partial I}{\partial v} = \frac{\partial I}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial I}{\partial y} \frac{\partial y}{\partial v} \quad (3.15)$$

Les composantes locales (g, t) s'expriment de la façon suivante :

$$\begin{cases} g = x \cos \phi + y \sin \phi \\ t = -x \sin \phi + y \cos \phi \end{cases} \quad (3.16)$$

Et la transformation inverse :

$$\begin{cases} x = g \cos \phi + t \sin \phi \\ y = g \sin \phi - t \cos \phi \end{cases} \quad (3.17)$$

Avec $\phi = \arg(\nabla I) = \arctan(I_y/I_x)$. Les dérivées d'ordre 1 à 3 selon les composantes (g, t) , valent donc :

$$\begin{cases} I_g = I_x \cos \phi + I_y \sin \phi = \|\nabla I\| \\ I_t = I_x \sin \phi - I_y \cos \phi = 0 \end{cases} \quad (3.18)$$

$$\begin{cases} I_{gg} = I_{xx} \cos^2 \phi + 2I_{xy} \cos \phi \sin \phi + I_{yy} \sin^2 \phi \\ I_{tt} = I_{xx} \sin^2 \phi - 2I_{xy} \cos \phi \sin \phi + I_{yy} \cos^2 \phi \end{cases} \quad (3.19)$$

$$\begin{cases} I_{ggg} = I_{xxx} \cos^3 \phi + 3I_{xxy} \cos^2 \phi \sin \phi + 3I_{xyy} \cos \phi \sin^2 \phi + I_{yyy} \sin^3 \phi \\ I_{ttt} = I_{xxx} \sin^3 \phi - 3I_{xxy} \cos \phi \sin^2 \phi - 3I_{xyy} \cos^2 \phi \sin \phi + I_{yyy} \cos^3 \phi \end{cases} \quad (3.20)$$

Formellement, les contours peuvent être définis comme le lieu des points où la dérivée seconde selon g s'annule, et la dérivée troisième selon g est strictement négative :

$$\begin{cases} I_{gg} = 0 \\ I_{ggg} < 0 \end{cases} \quad (3.21)$$

En remplaçant $\cos \phi$ et $\sin \phi$ par leur expression en fonction des coordonnées du gradient :

$$\begin{cases} I_{gg} = \frac{I_{xx}I_x^2 + 2I_{xy}I_xI_y + I_{yy}I_y^2}{I_x^2 + I_y^2} = 0 \\ I_{ggg} = \frac{I_{xxx}I_x^3 + 3I_{xxy}I_x^2I_y + 3I_{xyy}I_xI_y^2 + I_{yyy}I_y^3}{I_x^2 + I_y^2} < 0 \end{cases} \quad (3.22)$$

Comme on ne s'intéresse qu'au signe, on peut exclure le dénominateur, et la définition des contours devient finalement :

$$\begin{cases} I_{xx}I_x^2 + 2I_{xy}I_xI_y + I_{yy}I_y^2 = 0 \\ I_{xxx}I_x^3 + 3I_{xxy}I_x^2I_y + 3I_{xyy}I_xI_y^2 + I_{yyy}I_y^3 < 0 \end{cases} \quad (3.23)$$

L'implantation discrète consiste à calculer, en tout point (i, j) les deux quantités précédentes, soit :

$$\begin{cases} \hat{I}_{gg} = I_{xx}I_x^2 + 2I_{xy}I_xI_y + I_{yy}I_y^2 \\ \hat{I}_{ggg} = I_{xxx}I_x^3 + 3I_{xxy}I_x^2I_y + 3I_{xyy}I_xI_y^2 + I_{yyy}I_y^3 \end{cases} \quad (3.24)$$

Et à détecter les changements de polarité de \hat{I}_{gg} sur des voisinages 2×2 , sous condition de négativité de \hat{I}_{ggg} , soit :

$$\mathcal{C}_1[i, j] = 1 \Leftrightarrow \begin{cases} \bigvee_{(di, dj) \in \{(0,1), (1,0), (1,1)\}} (\hat{I}_{gg}[i, j] \cdot \hat{I}_{gg}[i - di, j - dj] < 0) \\ \bigwedge_{(di, dj) \in \{0,1\}^2} (\hat{I}_{ggg}[i - di, j - dj] < 0) \end{cases} \quad (3.25)$$

Avec \vee (resp. \wedge), le OU (resp. le ET) logique. La figure 3.10 montre le résultat de cet opérateur sur une image naturelle. On voit que l'application de la seule condition ($I_{gg} = 0; I_{ggg} < 0$) fournit beaucoup de contours, en effet, aucune condition sur la valeur du contraste n'est spécifiée, et de nombreuses structures imperceptibles sur l'image originale apparaissent. Pour limiter ce problème, on ajoute une condition en imposant, en plus, que le module du gradient soit supérieur à un certain seuil. Cependant, il est très difficile de trouver avec un seul seuil le bon compromis entre élimination des structures parasites et bonne continuité des contours intéressants (voir Figure 3.10, images (2) et (3)).

Pour cela, on utilise un seuillage par hystérésis, qui consiste à calculer 2 seuils, un seuil haut s_h et un seuil bas s_b . Chaque seuil fournit un ensemble de pixels, certains étant connectés entre eux pour former des courbes connexes. Le principe est de ne retenir que les courbes connexes de s_b qui ont une intersection non vide avec s_h . Cet opérateur correspond à la reconstruction géodésique de s_h dans s_b (voir Cours de Morphologie Mathématique).

L'opérateur \mathcal{C}_1 , bien que conceptuellement simple, est relativement complexe du point de vue calculatoire, car il implique le calcul de 9 grandeurs différentielles. Une approximation nettement moins coûteuse peut être obtenue en négligeant la condition de négativité de la dérivée troisième et en remplaçant les passages par zéro de la dérivée seconde dans la direction du gradient par les passages par zéros du laplacien. En observant l'équation 3.19, on peut remarquer que :

$$I_{gg} + I_{tt} = I_{xx} + I_{yy} \quad (3.26)$$

Le laplacien est donc un invariant vis-à-vis de la rotation. D'autre part, la dérivée seconde dans la direction tangentielle I_{tt} , est proportionnelle à la courbure de la ligne de niveau. Si l'on suppose que cette courbure est suffisamment faible, le laplacien coïncide alors avec la dérivée seconde dans la direction du gradient :

$$I_{tt} \simeq 0 \Rightarrow I_{gg} \simeq I_{xx} + I_{yy} \quad (3.27)$$

La définition formelle des contours devient donc :

$$\Delta_I = 0 \quad (3.28)$$

où $\Delta_I = I_{xx} + I_{yy}$ désigne le laplacien de I . Notons qu'il s'agit du détecteur de contours proposé par D. Marr et E. Hildreth [32]. Comme dans le cas précédent, on calcule le laplacien en tout point (i, j) , puis on détecte les changements de polarité de Δ_I sur des voisinages 2×2 :

$$\mathcal{C}_2[i, j] = 1 \Leftrightarrow \bigvee_{(di, dj) \in \{(0,1), (1,0), (1,1)\}} (\Delta_I[i, j] \cdot \Delta_I[i - di, j - dj] < 0) \quad (3.29)$$

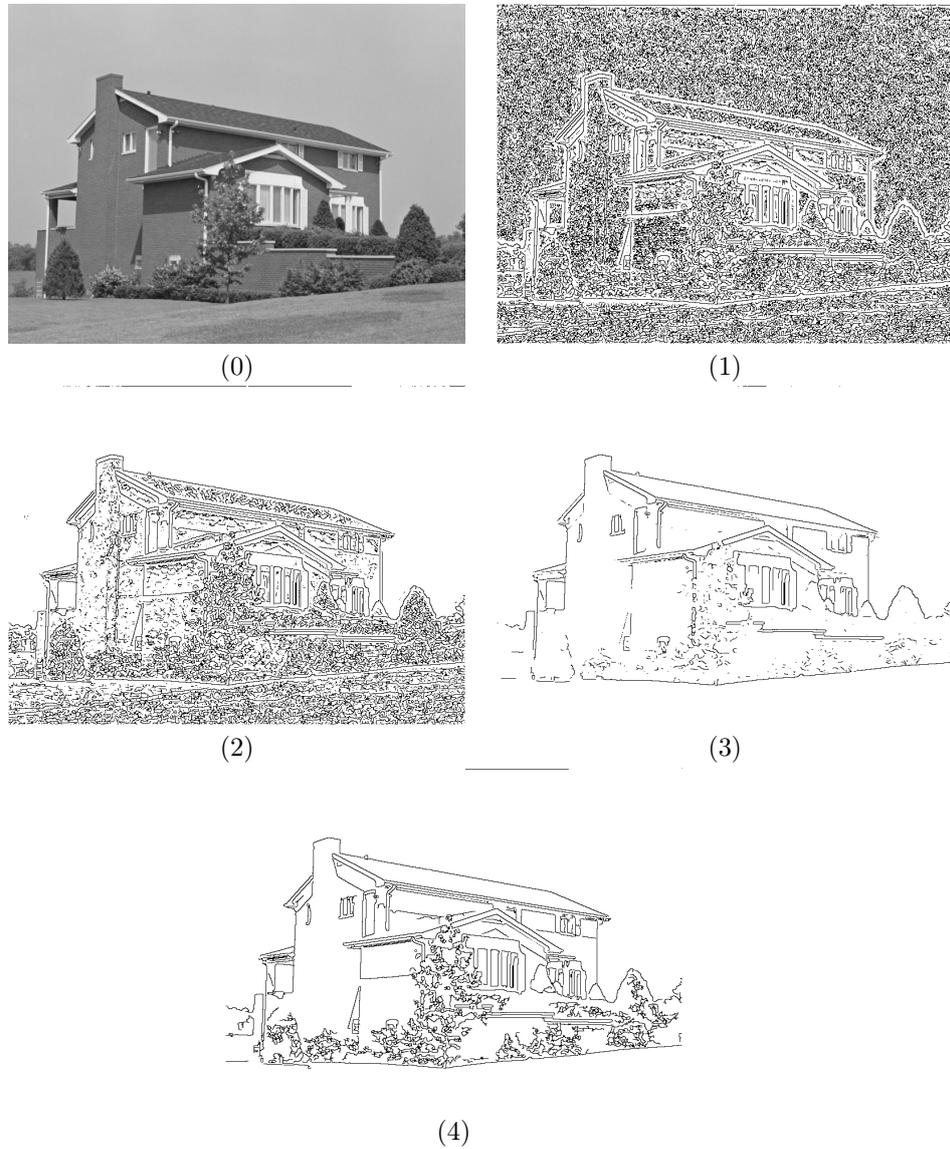


FIGURE 3.10 – Détection de contours par l'opérateur $I_{gg} = 0; I_{ggg} < 0$, pour l'échelle $s = 1.125$. (0) Image originale, (1) Détection sans seuil, (2) Détection avec seuil $\|\nabla I\| > 1.0$, (3) Détection avec seuil $\|\nabla I\| > 10.0$, (4) Détection avec seuillage par hystérésis (seuil bas : 1.0, seuil haut : 10.0)

La figure 3.11 montre le résultats de cet opérateur, dans les mêmes conditions que l'opérateur précédent.

Les figures 3.12 et 3.13 montrent les résultats obtenus par les opérateurs \mathcal{C}_1 , et \mathcal{C}_2 , pour différentes échelles.

3.5 Points d'intérêt

Les points d'intérêt constituent, au même titre que les contours, un opérateur de réduction d'information très important en traitement d'images. L'objectif est de trouver un ensemble de points isolés dans l'image, tel que chaque point soit représentatif de l'information locale, et que cette information soit suffisamment riche pour être discriminée du reste de l'image. Sont donc exclus les points appartenant à des régions homogènes, mais aussi les points qui se trouvent le long d'un contour rectiligne ou de faible courbure, car dans ce cas, les voisinages de tous les points qui se trouvent le long de ce contour sont semblables.

Les points d'intérêt sont donc associés à des points anguleux, aux voisinages desquels l'image varie significativement dans plusieurs directions.

Soit (x, y) un pixel. Soit $W_{(x,y)}$ une fenêtre centrée en (x, y) , représentant un voisinage du pixel. Une mesure des variations locales de l'image I au point (x, y) , et associée au déplacement $(\delta x, \delta y)$, est fournie par la fonction d'autocorrélation χ , définie comme suit :

$$\chi_{(x,y)}(\delta x, \delta y) = \sum_{(i,j) \in W_{(x,y)}} (I(i, j) - I(i + \delta x, j + \delta y))^2 \quad (3.30)$$

La formule de Taylor appliquée au premier ordre donne :

$$I(i + \delta x, j + \delta y) \simeq I(i, j) + \begin{pmatrix} \frac{\partial I}{\partial x}(i, j) & \frac{\partial I}{\partial y}(i, j) \end{pmatrix} \cdot \begin{pmatrix} \delta x \\ \delta y \end{pmatrix} \quad (3.31)$$

Et donc :

$$\chi_{(x,y)}(\delta x, \delta y) = \sum_{(i,j) \in W_{(x,y)}} \left(\begin{pmatrix} \frac{\partial I}{\partial x}(i, j) & \frac{\partial I}{\partial y}(i, j) \end{pmatrix} \cdot \begin{pmatrix} \delta x \\ \delta y \end{pmatrix} \right)^2 \quad (3.32)$$

Ce qu'on peut réécrire sous la forme matricielle suivante :

$$\chi_{(x,y)}(\delta x, \delta y) = \begin{pmatrix} \delta x & \delta y \end{pmatrix} \Xi_I(x, y) \begin{pmatrix} \delta x \\ \delta y \end{pmatrix} \quad (3.33)$$

Où $\Xi_I(x, y)$ est dite *matrice d'autocorrélation* de l'image I en (x, y) , avec :

$$\Xi_I(x, y) = \begin{pmatrix} \sum_{(i,j) \in W_{(x,y)}} \left(\frac{\partial I}{\partial x}(i, j) \right)^2 & \sum_{(i,j) \in W_{(x,y)}} \left(\frac{\partial I}{\partial x}(i, j) \cdot \frac{\partial I}{\partial y}(i, j) \right) \\ \sum_{(i,j) \in W_{(x,y)}} \left(\frac{\partial I}{\partial x}(i, j) \cdot \frac{\partial I}{\partial y}(i, j) \right) & \sum_{(i,j) \in W_{(x,y)}} \left(\frac{\partial I}{\partial y}(i, j) \right)^2 \end{pmatrix} \quad (3.34)$$

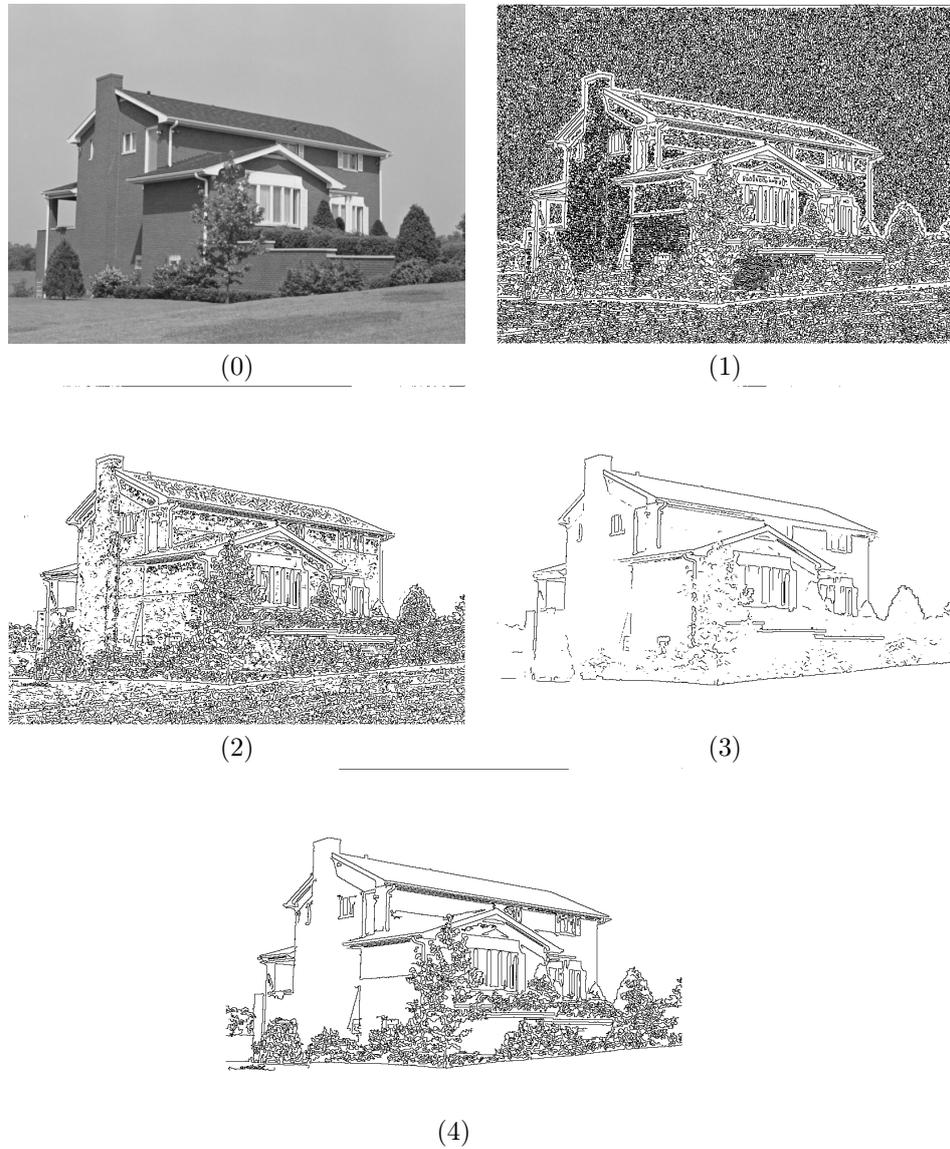


FIGURE 3.11 – Détection de contours par l'opérateur $\Delta_I = 0$, pour l'échelle $s = 1.125$. (0) Image originale, (1) Détection sans seuil, (2) Détection avec seuil $\|\nabla I\| > 1.0$, (3) Détection avec seuil $\|\nabla I\| > 10.0$, (4) Détection avec seuillage par hystérésis (seuil bas : 1.0, seuil haut : 10.0)

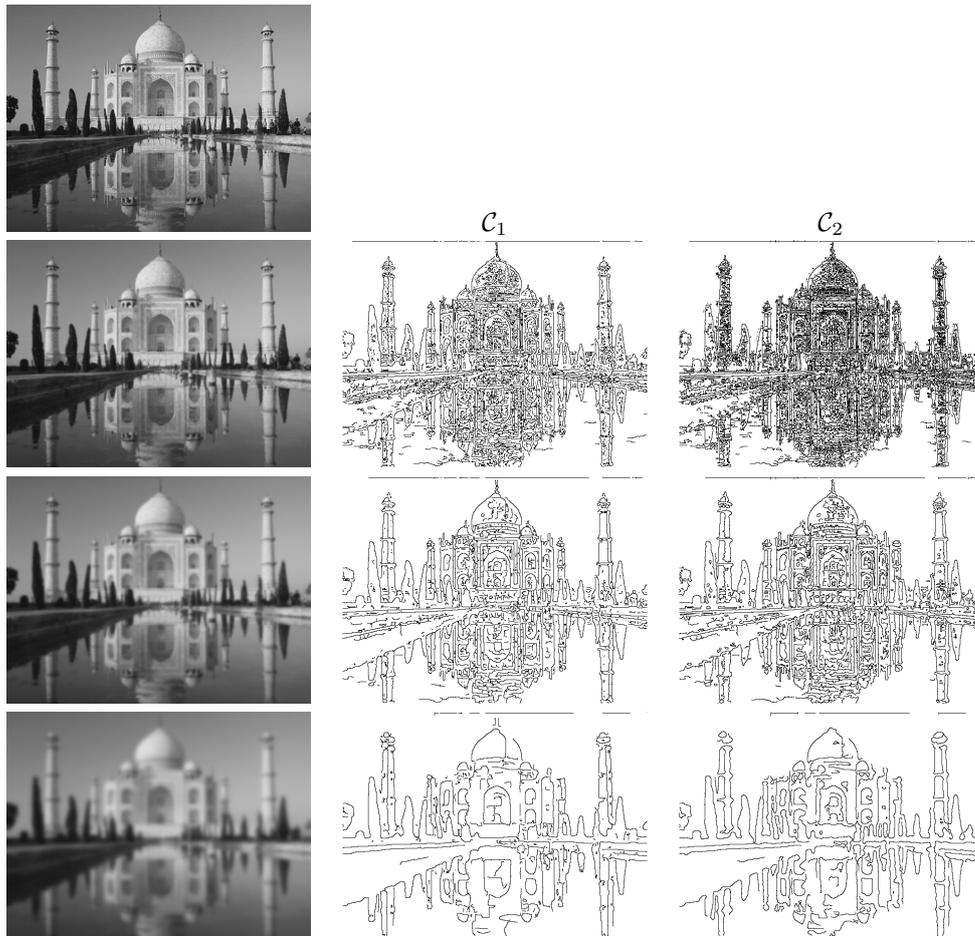


FIGURE 3.12 – Detections de contours et espaces d'échelles, pour les échelles $s = 0.5, 2, 8$. La colonne de gauche montre l'image lissée par la gaussienne de variance $2s$. La colonne du milieu montre les contours définis par $(I_{gg} = 0; I_{ggg} < 0)$, la colonne de droite les contours définis par $\Delta_I = 0$. Le même seuillage par hystérésis du gradient (seuil bas : 1.0; seuil haut : 2.0) a été appliqué pour toutes les échelles.

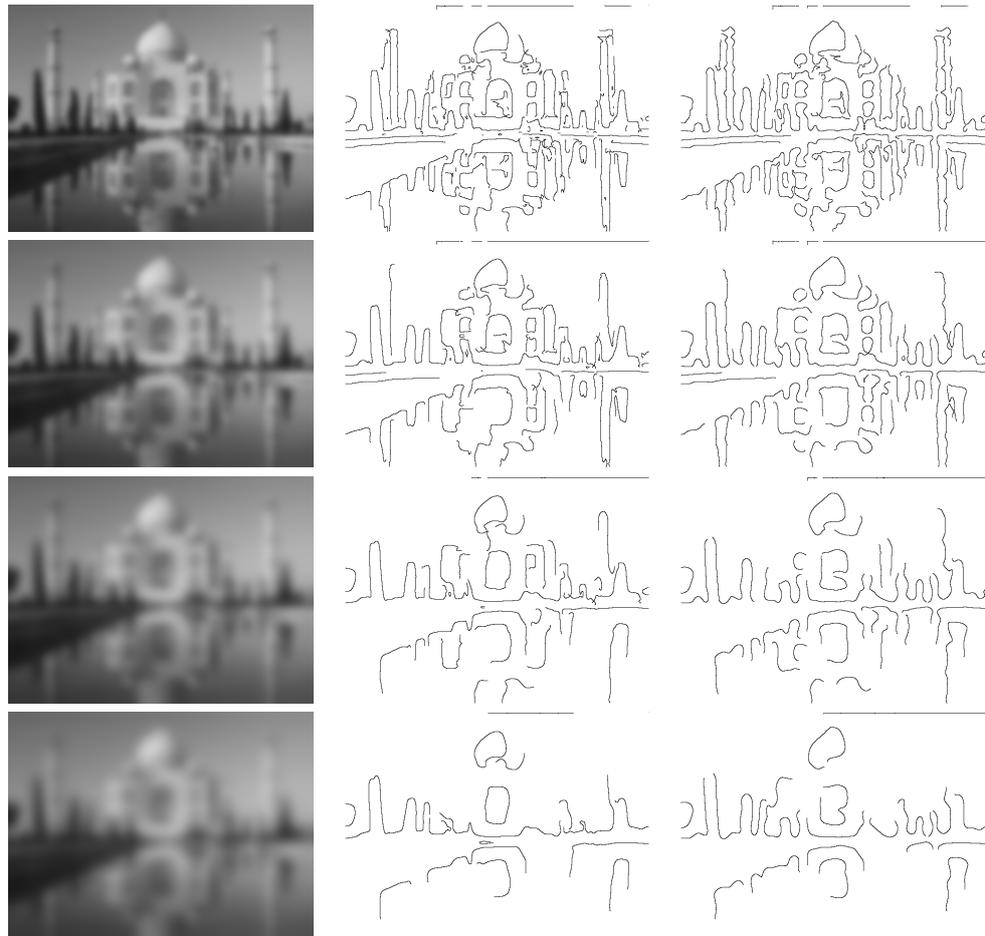


FIGURE 3.13 – Détection de contours et espaces d'échelles, pour les échelles $s = 18, 32, 50, 72$. La colonne de gauche montre l'image lissée par la gaussienne de variance $2s$. La colonne du milieu montre les contours définis par $(I_{gg} = 0; I_{ggg} < 0)$, la colonne de droite les contours définis par $\Delta_I = 0$. Le même seuillage par hystérésis du gradient (seuil bas : 1.0 ; seuil haut : 2.0) a été appliqué pour toutes les échelles.

La matrice d'autocorrélation $\Xi_I(x, y)$ représente la variation locale de l'image I au point (x, y) . Le pixel (x, y) sera considéré comme un point d'intérêt si, pour tous les déplacements $(\delta x, \delta y)$, la quantité $(\delta x, \delta y)\Xi_I(x, y)(\delta x, \delta y)^t$ est grande.

Les points d'intérêt sont donc les points (x, y) tels que la matrice d'autocorrélation $\Xi_I(x, y)$ a 2 valeurs propres grandes. Cela correspond aux points pour lesquels il existe localement une base de vecteurs propres de Ξ_I décrivant les principales directions de variations de l'image.

Le *détecteur de Harris* [16] est un opérateur basé sur la matrice d'autocorrélation. Il ne calcule pas explicitement les valeurs propres de Ξ_I , mais une fonction d'intérêt Θ_I définie par une combinaison linéaire du déterminant et de la trace de Ξ_I :

$$\Theta_I(x, y) = \det \Xi_I(x, y) - \alpha \operatorname{trace}^2 \Xi_I(x, y) \quad (3.35)$$

Le premier terme correspond au produit des valeurs propres, le second terme sert à pénaliser les points de fort contraste avec une seule forte valeur propre. α est un paramètre de pondération déterminé de façon empirique (les valeurs rencontrées dans la littérature varient peu, entre 0.04 et 0.06).

Les figures 3.14 et 3.15 montrent le calcul à différentes échelles de la fonction d'intérêt, pour une image faisant apparaître différents types de courbures. Noter l'effet du terme négatif faisant intervenir la trace, qui fait apparaître des valeurs négatives sur les contours rectilignes. Pour calculer la fonction d'intérêt de Harris Θ_I^s à l'échelle s , on calcule les dérivées premières I_x^s et I_y^s comme les convolutions de I avec les dérivées de gaussienne de variance $2s$, selon x et selon y respectivement. Puis, on calcule les composantes de la matrice Ξ_I^s en remplaçant la somme des dérivées dans la fenêtre W par un lissage gaussien de variance $2s'$. s et s' doivent être proportionnels, typiquement $s' = 2s$.

Une fois la fonction d'intérêt calculée, les points d'intérêt sont définis comme les maxima locaux de la fonction d'intérêt (c'est-à-dire les pixels qui ont une valeur de Θ_I supérieure ou égale à celle de tous leurs voisins. On fixe aussi habituellement un seuil sur la valeur de la fonction d'intérêt (par exemple 1% de la valeur maximale sur toute l'image de la fonction d'intérêt). La figure 3.16 montre les points d'intérêt détectés sur un détail de l'image «Guernica», pour différentes échelles.

3.6 Espaces d'échelles non linéaires

Les espaces d'échelles sont liés aux équations aux dérivées partielles (EDP) par la relation fondamentale entre la convolution gaussienne et l'équation de la chaleur, qui fournit la loi d'évolution d'un champ scalaire u (la température) au cours du temps t , dans un espace (le matériau) de conductivité c constante :

$$\frac{\partial u}{\partial t} = \operatorname{div}(c\nabla u) = c\Delta u, \quad (3.36)$$

avec div l'opérateur de divergence, défini en 2d par $\operatorname{div}(v) = \frac{\partial v}{\partial x} + \frac{\partial v}{\partial y}$. La solution de cette équation avec un champ initial u_0 est donnée par $u_t = u_0 \star G_t$, où G_t est le noyau gaussien d'écart-type $\sqrt{2t}$. Dans cette modélisation, notre image est associée au champ scalaire $I = u$, et il y a donc identité entre le facteur de temps et d'échelle. Ceci permet de voir les espaces d'échelles linéaires



FIGURE 3.14 – Fonction d'intérêt de Harris calculée sur l'image «Guernica», à l'échelle $s = 0.5$

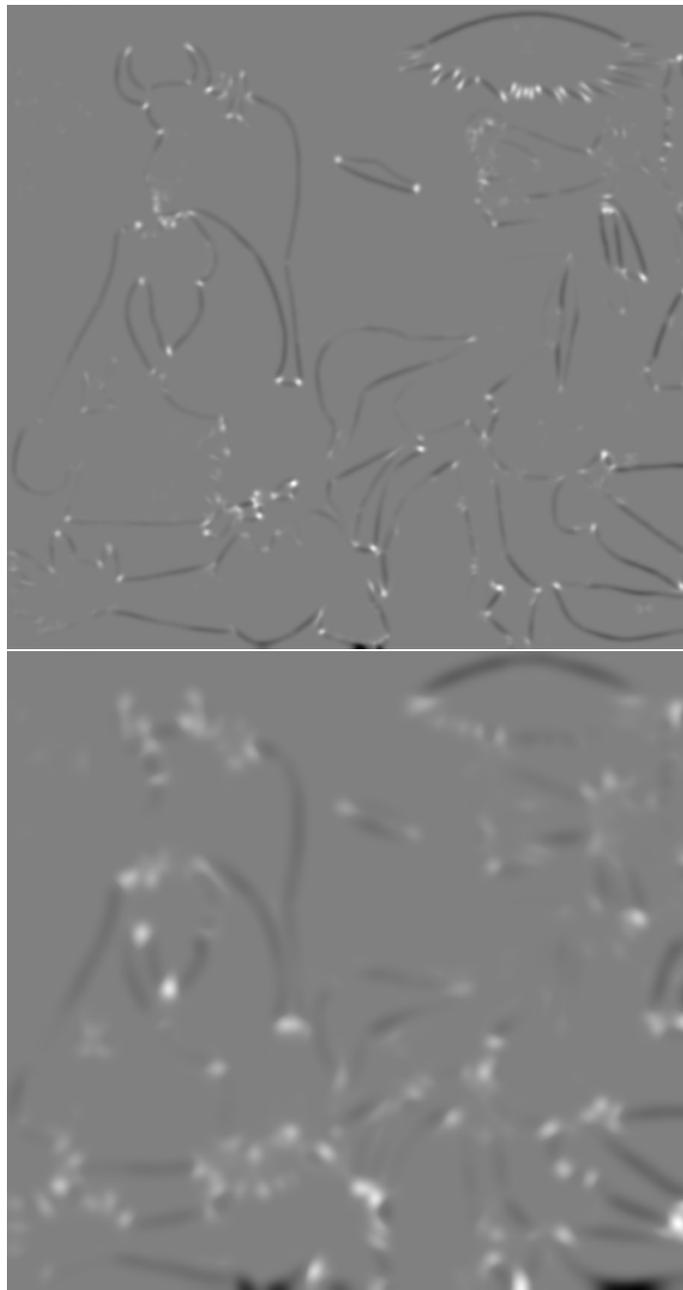


FIGURE 3.15 – Fonction d'intérêt de Harris calculée sur l'image «Guernica», aux échelles $s = 4.5$, et $s = 32.0$.

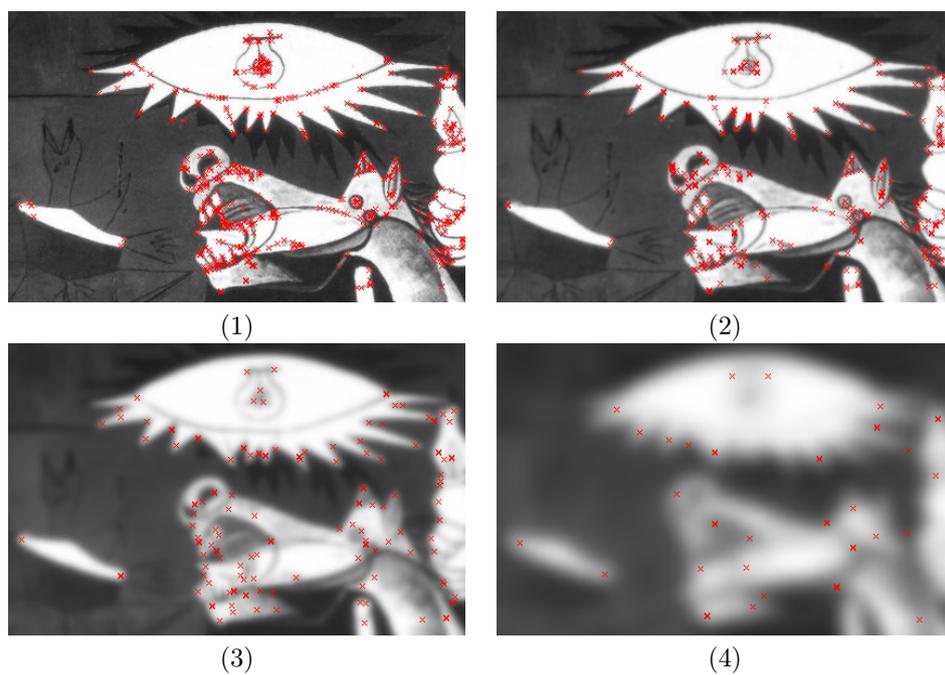


FIGURE 3.16 – Détection de points d'intérêt par la méthode de Harris multi-échelle, pour les échelles : (1) $s = 0.125$, (2) $s = 0.5$, (3) $s = 4.5$, et (4) $s = 32.0$. Les points affichés (croix en rouge) correspondent aux maxima locaux de la fonction d'intérêt Θ_I , dont la valeur de est supérieure à 1% du maximum (global) de Θ_I

comme un phénomène de diffusion isotrope (les valeurs se diffusent dans toutes les directions de l'espace à la même vitesse).

Mais l'intérêt principal de la modélisation par EDP est de généraliser le principe des espaces d'échelles, en transformant la conductivité c en un champ C variable dans l'espace :

$$\frac{\partial I}{\partial t} = \operatorname{div}(C\nabla I) = C\Delta I + \nabla C \cdot \nabla I \quad (3.37)$$

C'est ce qui forme la base des espaces d'échelles non linéaires. Au même titre que les filtres non linéaires, les espaces d'échelles non linéaires visent à simplifier progressivement l'image, tout en préservant les structures les plus importantes (typiquement les principaux contours). En pratique on définit donc la conductivité C en fonction du contenu de l'image, par exemple $C = g(\|\nabla I\|)$, où g est une fonction réelle décroissante, pour que les valeurs se diffusent préférentiellement dans les directions où l'image varie moins.

Néanmoins, contrairement au cas linéaire, on ne connaît pas la solution analytique en général, et la diffusion à une échelle donnée est donc calculée de façon itérative par un schéma de résolution numérique.

Ainsi, le schéma de diffusion anisotropique de Malik et Perona [36] est défini par l'équation d'évolution discrète suivante :

$$I^{(t+1)}(\mathbf{p}) = I^{(t)}(\mathbf{p}) + \lambda \sum_{\mathbf{q} \in \mathcal{N}(\mathbf{p})} (I^{(t)}(\mathbf{q}) - I^{(t)}(\mathbf{p})) \cdot g(|I^{(t)}(\mathbf{q}) - I^{(t)}(\mathbf{p})|) \quad (3.38)$$

avec $\mathcal{N}(\mathbf{p})$ l'ensemble des voisins du pixel \mathbf{p} , g une fonction réelle décroissante, et λ le paramètre de vitesse du schéma numérique, tel que $0 < \lambda < \operatorname{card}(\mathcal{N}(\mathbf{p}))$.

La fonction g peut être de type exponentielle : $g(x) = e^{-\left(\frac{x}{k}\right)^2}$, ou de type hyperbolique : $g(x) = \frac{1}{1+\left(\frac{x}{k}\right)^2}$. Dans les deux cas, k est un paramètre de dégénérescence du filtre, qui dépend du niveau de régularisation désiré, c'est-à-dire du niveau de réduction des différences locales, indépendamment de l'échelle. On peut remarquer que lorsque k tend vers l'infini, la partie incrémentale de la loi d'évolution devient

$$I^{(t+1)}(\mathbf{p}) - I^{(t)}(\mathbf{p}) = \lambda \left(-I^{(t)}(\mathbf{p})\operatorname{card}(\mathcal{N}(\mathbf{p})) + \sum_{\mathbf{q} \in \mathcal{N}(\mathbf{p})} I^{(t)}(\mathbf{q}) \right).$$

On peut reconnaître dans la grande parenthèse du terme de droite la discrétisation sur $\mathcal{N}(\mathbf{p})$ du laplacien (voir Chapitre 2), ce qui correspond au schéma discret de la diffusion isotrope, donc à l'implantation itérative de la convolution gaussienne.

Dans le cas non dégénéré, le schéma de diffusion anisotropique produit un lissage progressif des images qui conserve les gradients les plus élevés. Voir un exemple Figure 3.17 sur 4 échelles.



FIGURE 3.17 – Espace d'échelle non linéaire de diffusion anisotropique, de Malik et Perona, pour 4 échelles t différentes, avec $k = 25$, $\alpha = 0.1$, et schéma de décroissance hyperbolique.